Universal Scheduling for Networks with Arbitrary Traffic, Channels, and Mobility

Michael J. Neely

Abstract—We extend stochastic network optimization theory to treat networks with arbitrary sample paths for arrivals, channels, and mobility. The network can experience unexpected link or node failures, traffic bursts, and topology changes, and there are no probabilistic assumptions describing these time varying events. Performance of our scheduling algorithm is compared against an ideal T-slot lookahead policy that can make optimal decisions based on knowledge up to T-slots into the future. We develop a simple non-anticipating algorithm that provides network throughput-utility that is arbitrarily close to (or better than) that of the T-slot lookahead policy, with a tradeoff in the worst case queue backlog kept at any queue. The same policy offers even stronger performance, closely matching that of an ideal *infinite lookahead* policy, when ergodic assumptions are imposed. Our analysis uses a sample path version of Lyapunov drift and provides a methodology for optimizing time averages in general time-varying optimization problems.

Index Terms—Queueing analysis, opportunistic scheduling, internet, routing, flow control, wireless networks, optimization

I. INTRODUCTION

Networks experience unexpected events. Consider the network of Fig. 1 and focus on the session that sends a stream of packets from node A to node D. Suppose that several paths are used, but due to congestion on other links, the primary path that can deliver the most data is the path A, B, C, D. However, suppose that there is a failure at node B in the middle of the session. An algorithm with perfect knowledge of the future would take advantage of the path A, B, C, D while it is available, and would switch to alternate paths before the failure occurs. The algorithm would also be able to predict the traffic load on different links at different times, and would optimally route in anticipation of these events.

The above example holds if the network of Fig. 1 is a wireline network, a wireless network, or a mixture of wired and wireless connections. As another example, suppose the network contains an additional mobile wireless node E, and that the following unexpected event occurs: Node E moves into close proximity to node A, allowing a large number of packets to be sent to it. It then moves into close proximity to node D, providing an opportunity to transmit packets to this destination node. If this event could be anticipated, we could
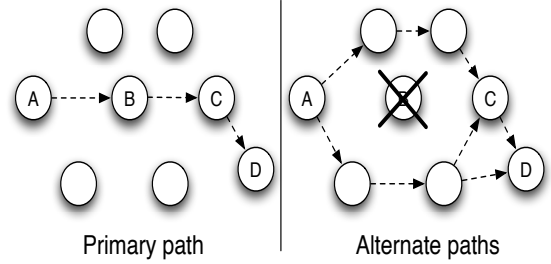
Fig. 1. A primary path from A to D, with alternative paths shown in the event of a failure at node B.

take advantage of it and improve the short term throughput by routing many packets over the relay E.

These examples illustrate different types of unexpected events that can be exploited to improve performance. There are of course even more complex sequences of arrival, channel, and mobility events that, if known in advance, could be exploited to yield improved performance. However, because realistic networks do not have knowledge of the future, it is not clear if these events can be practically used. Surprisingly, this paper shows that it is possible to reap the benefits of these time varying events without any knowledge of the future. We show that a simple non-anticipating policy can closely track the performance of an ideal T-*slot lookahead policy* that has perfect knowledge of the future up to T slots. Proximity to the performance of the T-slot lookahead policy comes with a corresponding tradeoff in the worst case queue backlog stored in any queue of the network, which also affects a tradeoff in network delay.

Specifically, we treat networks with slotted time with normalized slots $t \in \{0, 1, 2, \ldots\}$. We measure network utility over an interval of timeslots according to a concave function of the time average throughput vector achieved over that interval. We show that for any positive integer frame size T, and any interval that consists of R frames of T slots, the utility achieved over the interval is greater than or equal to the utility achieved by using the T-slot lookahead policy over each of the R frames, minus a "fudge factor" that has the form:

$$fudge\ factor = \frac{B_1 T}{V} + \frac{B_2 V}{RT}$$

where B_1 and B_2 are constants, and V is a positive parameter that can be chosen as desired to make the term $B_1 T/V$ arbitrarily small, with a tradeoff in the worst case queue backlog that is $O(V)$. This shows that we reap almost the same benefits of knowing the future up to T slots if we

choose V suitably large and if we wait for the completion of R frames of size T, where R is sufficiently large to make $B_2V/(RT)$ small. Remarkably, the constants B_1 and B_2 can be explicitly computed in advance, without any assumptions on the underlying stochastic processes that describe the time varying events.

This establishes a *universal scheduling paradigm* that shows a single network algorithm can provide strong mathematical guarantees for any network and for any time varying sample paths. The algorithm that we use is not new: It is a modified version of the backpressure based "drift-minus-reward" algorithms that we previously developed and used in different contexts in our prior work [1][2][3][4]. These algorithms were originally developed for the case when new arrivals and new channel states are independent and identically distributed (i.i.d.) over slots, and were analyzed using a Lyapunov drift defined as an expectation over the underlying probability distribution. However, it is known that Lyapunov based algorithms that are designed under i.i.d. assumptions yield similar performance under more general ergodic (but non-i.i.d.) probability models. This is shown for network stability using a T-slot Lyapunov drift in [5][6], and using a related delayed-queue analysis (that often provides tighter delay bounds) in [7]. Further, such algorithms are known to be robust to non-ergodic situations such as when traffic yields "instantaneous rates" that can vary arbitrarily inside the capacity region [8][4][1], and when "instantaneous capacity regions" can vary arbitrarily but are assumed to always contain the traffic rate vector [9]. However, the prior non-ergodic analysis [8][4][1][9] still assumes an underlying probability model, and makes assumptions about traffic rates and network capacity with respect to this model.

The analysis in this paper is new and uses a sample path version of Lyapunov drift, without any probabilistic assumptions. This framework allows treatment of realistic channels and traffic traces. Because arbitrary sample paths may not have well defined time averages, typical *equilibrium* notions of *network capacity* and *optimal time average utility* cannot be used. We thus use a new metric that measures performance with respect to ideal T-slot lookahead policies. This is a possible framework for treating the important open questions identified in [10] concerning *non-equilibrium network theory*. Further, our results provide universal techniques for optimizing time averages that are useful for other types of time-varying systems.

A. Comparison to Related Work

We note that *universal algorithms* are important in other fields. For example, the universal Lempel-Ziv data compression algorithm operates on arbitrary files [11], and universal stock portfolio allocation algorithms hold for arbitrary price sample paths [12][13][14][15]. Our work provides a universal approach to network scheduling. It is important to note that prior work in the area of *competitive ratio analysis* [16][17][18][19] and *adversarial queueing theory* [20] also considers network scheduling problems with arbitrary sample paths, albeit in a different context. Work in [17] considers a

large class of admission control problems for networks with random arrivals that earn revenue if accepted. An algorithm is developed that yields revenue that differs by a factor of $\Theta(\log(N))$ from that of an ideal algorithm with perfect knowledge of the future, where N is the number of network nodes. Further, this asymptotic ratio is shown to be optimal, in the sense that there is always a *worst case* sequence of packet arrivals that can reduce revenue by this amount. Related $\Theta(\log(N))$ competitive ratio results are developed for energy optimization in [18] and for wireless admission control in [19]. The works [16][17][18][19] do not consider networks with time varying channels or mobility, and do not treat (or exploit) network queueing. An adversarial queueing theory example in [20] shows that, if channels are time varying, the competitive ratio can be much worse than logarithmic, even for a simple packet-based network with a single link.

Our work treats the difficult case of multi-hop networks with arbitrary time varying channels, mobility, and penalty constraints. However, rather than pursuing a competitive ratio analysis, we measure performance against a T-slot lookahead metric. We show that we can closely track the performance of an ideal T-slot lookahead policy, for any arbitrary (but finite) T. This does not imply that the algorithm has an optimal competitive ratio, because the utility of a T-slot lookahead policy for finite T may not be as good as the performance of an *infinite* lookahead policy. However, it turns out that our policy indeed approaches an optimal competitive ratio (measured with respect to an infinite lookahead policy) under the special case when the time varying events are ergodic.

Finally, we note that a frame-based metric, similar to our T-slot lookahead metric, was used in [20] to treat static wireline networks with arbitrary arrivals but fixed topology and channel states. There, an algorithm that queues all packets that arrive in a frame, solves a network-wide utility maximization problem for these packets based on knowledge of the static link capacities, and implements the solution in the next frame, is shown to achieve revenue that is close to that of a policy that a-priori knows the packet arrival times over one frame. In our context, we do not have the luxury of solving a network-wide utility maximization problem based on known static link capacities, because the network itself is changing with time. Our solution strategy is thus completely different from that of [20]. Rather than a frame-based approach, our algorithm makes simple "max-weight" decisions every slot based on a quadratic Lyapunov function. It is interesting to note that this imposes a "cost" associated with each decision that depends on the current network queue state, which is similar in spirit to the cost functions used in the algorithms of [16][17][18][19] for competitive ratio analysis.

B. Outline of Paper

The next section describes the general problem of constrained optimization of time varying systems. Section III provides a universal solution technique that measures performance against a T-slot lookahead policy. Section IV applies the framework to a simple internet model, and Section V applies the framework to a more extensive class of time varying networks.

II. GENERAL TIME VARYING OPTIMIZATION PROBLEMS

Here we provide a framework for universal constrained optimization for a general class of time varying systems. The framework is applied in Sections IV and V to solve the network problems of interest.

Consider a slotted system with normalized timeslots $t \in \{0, 1, 2, \ldots\}$. The system contains K queues with current backlog given by the vector $\boldsymbol{Q}(t) = (Q_1(t), \ldots, Q_K(t))$. Let $\omega(t)$ denote a *random event* that occurs on slot t. The random event $\omega(t)$ represents a collection of current system parameters and takes values in some abstract *event space* Ω. We treat $\omega(t)$ as a pre-determined function of $t \in \{0, 1, 2, \ldots\}$, although each value $\omega(t)$ is not revealed until the beginning of slot t. Every slot, a system controller observes the current value of $\omega(t)$ and chooses a *control action* $\alpha(t)$, constrained to some *action space* $\mathcal{A}_{\omega(t)}$ that can depend on $\omega(t)$. The random event $\omega(t)$ and the corresponding control action $\alpha(t) \in \mathcal{A}_{\omega(t)}$ produce a *service vector* $\boldsymbol{b}(t) = (b_1(t), \ldots, b_K(t))$, an *arrival vector* $\boldsymbol{a}(t) = (a_1(t), \ldots, a_K(t))$, and two *attribute vectors* $\boldsymbol{x}(t) = (x_1(t), \ldots, x_M(t))$, $\boldsymbol{y}(t) = (y_0(t), y_1(t), \ldots, y_L(t))$ (for some non-negative integers M and L). These vectors are general functions of $\omega(t)$ and $\alpha(t)$:

$$
\begin{aligned}
a_k(t) &= \hat{a}_k(\alpha(t), \omega(t)) \ \ \forall k \in \{1, \ldots, K\} \\
b_k(t) &= \hat{b}_k(\alpha(t), \omega(t)) \ \ \forall k \in \{1, \ldots, K\} \\
x_m(t) &= \hat{x}_m(\alpha(t), \omega(t)) \ \ \forall m \in \{1, \ldots, M\} \\
y_l(t) &= \hat{y}_l(\alpha(t), \omega(t)) \ \ \forall l \in \{0, 1, \ldots, L\}
\end{aligned}
$$

The queue dynamics are determined by the arrival and service variables by:

$$
Q_k(t+1) = \max[Q_k(t) - b_k(t), 0] + a_k(t) \tag{1}
$$

Let $\overline{x}_m$ be the time average of $x_m(t)$ under a particular control policy implemented over a finite number of slots t_{end}:

$$
\overline{x}_m \triangleq \frac{1}{t_{end}} \sum_{\tau=0}^{t_{end}-1} x_m(\tau) \tag{2}
$$

Define $\overline{a}_k$, $\overline{b}_k$, $\overline{y}_l$ similarly. Define $\overline{\boldsymbol{x}} \triangleq (\overline{x}_1, \ldots, \overline{x}_M)$. The goal is to design a policy that solves the following time average optimization problem:

$$
\begin{aligned}
\text{Minimize:} \quad & \overline{y}_0 + f(\overline{\boldsymbol{x}}) & (3) \\
\text{Subject to:} \quad & \overline{y}_l + g_l(\overline{\boldsymbol{x}}) \leq 0 \ \ \forall l \in \{1, \ldots, L\} & (4) \\
& \overline{a}_k \leq \overline{b}_k \ \ \forall k \in \{1, \ldots, K\} & (5) \\
& \overline{\boldsymbol{x}} \in \mathcal{X} & (6) \\
& \alpha(t) \in \mathcal{A}_{\omega(t)} \ \forall t \in \{0, \ldots, t_{end}-1\} & (7)
\end{aligned}
$$

where $f(\boldsymbol{x})$, $g_l(\boldsymbol{x})$ are convex *cost functions* of the vector $\boldsymbol{x} = (x_1, \ldots, x_M)$, and $\mathcal{X}$ is a general convex subset of $\mathbb{R}^M$. The above problem is of interest even if there are no underlying queues $Q_k(t)$ (so that the constraints (5) are removed), and/or if $\mathcal{X} = \mathbb{R}^M$ (so that the constraint (6) is removed), and/or if $f(\cdot) = g_l(\cdot) = 0$.

The above problem is stated in terms of a finite horizon of size t_{end}. The minimum cost in (3) is defined for a given $\omega(t)$ function over $t \in \{0, \ldots, t_{end}-1\}$, and considers all possible control actions that can be implemented over the time horizon, including actions that have full knowledge of the future values of $\omega(t)$. However, we desire a *non-anticipating* control policy that only knows the current $\omega(t)$ value on each slot t, and has no knowledge of the future. We note that a theory for solving stochastic network optimization problems similar to (3)-(7) (without the set constraint (6)) is developed in [1] for an infinite horizon context under the assumption that $\omega(t)$ is i.i.d. over slots with some (possibly unknown) probability distribution, and related problems are treated in a fluid limit sense in [21][22]. Here, we do not consider any probability model for $\omega(t)$, so that it is not possible to use law of large number averaging principles, or to achieve the optimum performance in an "expected" sense. Rather, we shall deterministically achieve an "approximate" optimum for large t_{end} values, to be made precise in future sections.

A. Boundedness and Feasibility Assumptions

Here we present assumptions concerning the functions $\hat{a}_k(\cdot)$, $\hat{b}_k(\cdot)$, $\hat{x}_m(\cdot)$, $\hat{y}_l(\cdot)$ that ensure the problem (3)-(7) is feasible with a bounded infimum cost.

Assumption A1: The functions $\hat{a}_k(\cdot)$, $\hat{b}_k(\cdot)$, $\hat{x}_m(\cdot)$, $\hat{y}_l(\cdot)$ are bounded, so that for all $\omega \in \{\omega(0), \ldots, \omega(t_{end} - 1)\}$, all $\alpha \in \mathcal{A}_\omega$, and all $k \in \{1, \ldots, K\}$, $m \in \{1, \ldots, M\}$, $l \in \{0, 1, \ldots, L\}$ we have:

$$
\begin{aligned}
0 &\leq \hat{a}_k(\alpha, \omega) \leq a_k^{max} \\
0 &\leq \hat{b}_k(\alpha, \omega) \leq b_k^{max} \\
x_m^{min} &\leq \hat{x}_m(\alpha, \omega) \leq x_m^{max} \\
y_l^{min} &\leq \hat{y}_l(\alpha, \omega) \leq y_l^{max}
\end{aligned}
$$

for some finite constants a_k^{max}, b_k^{max}, x_m^{min}, x_m^{max}, y_l^{min}, y_l^{max}. Further, the cost functions $f(\boldsymbol{x})$ and $g_l(\boldsymbol{x})$ are defined over all vectors $(x_1, \ldots, x_M)$ that satisfy $x_m^{min} \leq x_m \leq x_m^{max}$ for all $m \in \{1, \ldots, M\}$, and have finite upper and lower bounds f^{min}, f^{max}, g_l^{min}, g_l^{max} over this region.

Assumption A2: For all $\omega \in \{\omega(0), \ldots, \omega(t_{end}-1)\}$, there is at least one control action $\alpha'_\omega \in \mathcal{A}_\omega$ that satisfies:

$$
\begin{aligned}
\hat{y}_l(\alpha'_\omega, \omega) + g_l(\hat{\boldsymbol{x}}(\alpha'_\omega, \omega)) &\leq 0 \quad \forall l \in \{1, \ldots, L\} & (8) \\
\hat{a}_k(\alpha'_\omega, \omega) &\leq \hat{b}_k(\alpha'_\omega, \omega) \quad \forall k \in \{1, \ldots, K\} & (9) \\
\hat{\boldsymbol{x}}(\alpha'_\omega, \omega) &\in \mathcal{X} & (10)
\end{aligned}
$$

where $\hat{\boldsymbol{x}}(\alpha, \omega)$ is defined:

$$
\hat{\boldsymbol{x}}(\alpha, \omega) \triangleq (\hat{x}_1(\alpha, \omega), \ldots, \hat{x}_M(\alpha, \omega))
$$

For a given $\{\omega(0), \ldots, \omega(t_{end} - 1)\}$, we say that the problem (3)-(7) is *feasible* if there exist control actions $\{\alpha(0), \ldots, \alpha(t_{end} - 1)\}$ that satisfy the constraints (4)-(7). Assumption A2 ensures that the problem is feasible (just consider the control actions $\alpha(t) = \alpha'_{\omega(t)}$ for all t, and use Jensen's inequality to note that $g_l(\overline{\boldsymbol{x}}) \leq \overline{g_l(\boldsymbol{x})}$). Define F^* as the infimum value of the cost metric (3) over all feasible policies. The value F^* is finite by Assumption A1. Then for any $\epsilon > 0$, there are control actions $\{\alpha^*(0), \ldots, \alpha^*(t_{end}-1)\}$ that satisfy the constraints (4)-(7) with a total cost that satisfies:

$$
F^* \leq \overline{y}_0^* + f(\overline{\boldsymbol{x}}^*) \leq F^* + \epsilon
$$

Appendix B provides conditions that ensure the infimum F^* is achievable by a particular policy (i.e., with $\epsilon = 0$). We note that Assumption A2 can be relaxed to only require feasibility over frames of T slots, although the resulting performance bounds in Theorem 1a and 1b are slightly altered in this case.

B. Cost Function Assumptions

The cost functions $f(\boldsymbol{x})$, $g_l(\boldsymbol{x})$ are assumed to be convex and continuous over the region of all $(x_1, \ldots, x_M)$ vectors that satisfy:

$$x_m^{min} \leq x_m \leq x_m^{max} \quad \text{for all } m \in \{1, \ldots, M\} \quad (11)$$

In addition, we assume that the magnitude of the mth left and right partial derivatives of $f(\boldsymbol{x})$ with respect to x_m are upper bounded by finite constants $\nu_m \geq 0$ for all $\boldsymbol{x}$ that satisfy (11) and such that $x_m^{min} < x_m < x_m^{max}$.[1] Similarly, the magnitude of the right and left partial derivatives of $g_l(\boldsymbol{x})$ are upper bounded by finite constants $\beta_{l,m} \geq 0$. This implies that:

$$f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + \sum_{m=1}^{M} \nu_m |y_m| \quad (12)$$

$$g_l(\boldsymbol{x} + \boldsymbol{y}) \leq g_l(\boldsymbol{x}) + \sum_{m=1}^{M} \beta_{l,m} |y_m| \quad (13)$$

for all $\boldsymbol{x}$, $\boldsymbol{y}$ such that $\boldsymbol{x}$ and $\boldsymbol{x} + \boldsymbol{y}$ are in the region specified by (11).

An example of a non-differentiable cost function that satisfies all of the above assumptions is:

$$f(x_1, \ldots, x_M) = \max[x_1, \ldots, x_M]$$

In this case, we have $\nu_m = 1$ for all m. Another example is a separable cost function:

$$f(x_1, \ldots, x_M) = \sum_{m=1}^{M} f_m(x_m) \quad (14)$$

where functions $f_m(x)$ are continuous and convex with derivatives bounded in magnitude by ν_m over $x_m^{min} \leq x \leq x_m^{max}$.

C. Applications of This Optimization Framework

We show in Section V that this problem applies to general dynamic networks. There, the $\omega(t)$ value represents a collection of channel conditions for all network links on slot t. This includes the simple model where link conditions are either in the ON or OFF state, representing connections or disconnections that can vary from slot to slot due to fading channels and/or user mobility. Each node can discover the state of its links by probing to find existing neighbors on the current slot. The $\alpha(t)$ value represents a collection of routing, resource allocation, and/or flow control decisions that are taken by the network in reaction to the current $\omega(t)$ value.

In addition to dynamic networks and queueing systems, the problem (3) has applications in many other areas that involve optimization over time varying systems. For example, our recent work in [15] presents applications to stock market trading problems. There, $\omega(t)$ represents a vector of current stock prices, and the constraints $\overline{a}_k \leq \overline{b}_k$ ensure that the

---

[1] All convex functions have well defined right and left partial derivatives.

average amount of stock k sales cannot exceed the average amount of stock k purchases.

We note that Assumption A2, which assumes that for any ω there exists an action α'_ω that satisfies $\hat{a}_k(\alpha'_\omega, \omega) \leq \hat{b}_k(\alpha'_\omega, \omega)$, often holds for systems that have a physical "idle" control action that reduces the inequality to $0 \leq 0$. For example, in network problems, the values of $\hat{a}_k(\cdot)$ and $\hat{b}_k(\cdot)$ often represent transmission rates, power expenditures, or newly accepted jobs, and the idle action is the one that accepts no new arrivals into the network and transmits no data over any link of the network. For stock market problems, the idle action is often the action that neither buys nor sells any shares of stock on the current slot.

D. T-Slot Lookahead Policies

Rather than consider the optimum of the problem (3)-(7) over the full time interval $t \in \{0, \ldots, t_{end} - 1\}$, we consider the minimum cost that can be incurred over successive frames of size T, assuming that the time average constraints (4)-(6) must be achieved over each frame. Specifically, let T be a positive integer, representing a *frame size*. For a non-negative integer r, define F_r^* as the infimum value associated with the following problem (where $\boldsymbol{\gamma} \triangleq (\gamma_1, \ldots, \gamma_M)$):

Minimize: $\qquad h_0 + f(\boldsymbol{\gamma}) \qquad\qquad (15)$

Subject to: $\qquad h_l + g_l(\boldsymbol{\gamma}) \leq 0 \; \forall l \in \{1, \ldots, L\}$

$$\boldsymbol{\gamma} \in \mathcal{X}$$

$$\gamma_m = \tfrac{1}{T} \sum_{\tau=rT}^{(r+1)T-1} \hat{x}_m(\alpha(\tau), \omega(\tau)) \;\; \forall m \in \{1, \ldots, M\}$$

$$h_l = \tfrac{1}{T} \sum_{\tau=rT}^{(r+1)T-1} \hat{y}_l(\alpha(\tau), \omega(\tau)) \;\; \forall l \in \{0, 1, \ldots, L\}$$

$$\tfrac{1}{T} \sum_{\tau=rT}^{(r+1)T-1} [\hat{a}_k(\alpha(\tau), \omega(\tau)) - \hat{b}_k(\alpha(\tau), \omega(\tau))] \leq 0 \; \forall k$$

$$\alpha(\tau) \in \mathcal{A}_{\omega(\tau)} \;\; \forall \tau \in \{rT, \ldots, (r+1)T - 1\}$$

The value of F_r^* represents the infimum of the cost metric that can be achieved over the frame, considering all policies that satisfy the constraints and that have perfect knowledge of the future $\omega(\tau)$ values over the frame. Our new goal is to design a non-anticipating control policy that is implemented over time $t_{end} = RT$ (for some positive integer R), and that satisfies all constraints of the original problem while achieving a total cost that is close to (or smaller than) the value of:

$$\frac{1}{R} \sum_{r=0}^{R-1} F_r^* \quad (16)$$

For $R = 1$, it is clear that the value in (16) is the same as the optimal cost associated with the problem (3) with $t_{end} = T$. For $R \geq 1$, it can be shown that the value in (16) is greater than or equal to the optimal cost associated with the problem (3) for $t_{end} = RT$. The reason that the problem (3)-(7) might have a strictly smaller cost is that it only requires the time average constraints to be met over the full time interval, rather than requiring them to be satisfied on each of the R frames. Nevertheless, when T is large, it is not trivial to achieve the cost value of (16), as this cost is defined over policies that have T-slot lookahead, whereas an actual policy does not have future lookahead capabilities.

III. SOLUTION TO THE GENERAL PROBLEM

First note that the problem (3)-(7) is equivalent to the following problem, which introduces *auxiliary variables* $\boldsymbol{\gamma}(t) = (\gamma_1(t), \ldots, \gamma_M(t))$ for $t \in \{0, \ldots, t_{end} - 1\}$:

$$\text{Minimize:} \qquad \overline{y}_0 + \overline{f(\boldsymbol{\gamma})} \qquad (17)$$

$$\text{Subject to:} \qquad \overline{y}_l + \overline{g_l(\boldsymbol{\gamma})} \leq 0 \ \forall l \in \{1, \ldots, L\} \qquad (18)$$

$$\overline{a}_k \leq \overline{b}_k \ \forall k \in \{1, \ldots, K\} \qquad (19)$$

$$\overline{\gamma}_m = \overline{x}_m \ \forall m \in \{1, \ldots, M\} \qquad (20)$$

$$\alpha(t) \in \mathcal{A}_{\omega(t)} \ \forall t \in \{0, \ldots, t_{end} - 1\} \qquad (21)$$

$$\boldsymbol{\gamma}(t) \in \mathcal{X} \ \forall t \in \{0, \ldots, t_{end} - 1\} \qquad (22)$$

$$x_m^{min} \leq \gamma_m(t) \leq x_m^{max} \ \forall t \in \{0, \ldots, t_{end} - 1\} \qquad (23)$$

where $\overline{f(\boldsymbol{\gamma})}$ is defined:

$$\overline{f(\boldsymbol{\gamma})} \triangleq \frac{1}{t_{end}} \sum_{\tau=0}^{t_{end}-1} f(\boldsymbol{\gamma}(\tau))$$

and where $\overline{g_l(\boldsymbol{\gamma})}$ is defined similarly.

To see that the above problem (17)-(23) is equivalent to the original problem (3)-(7), note that any optimal solution of (3)-(7) also satisfies the constraints (18)-(23), with the same value of the cost metric (17), provided that we define $\gamma_m(t) = \overline{x}_m$ for all t, with $\overline{x}_m$ being the time average of $x_m(t)$ under the solution to the problem (3)-(7). Thus, the minimum cost metric of the new problem (17)-(23) is *less than or equal* to that of (3)-(7). On the other hand, by Jensen's inequality and convexity of $f(\boldsymbol{\gamma})$, $g_l(\boldsymbol{\gamma})$, we have for any solution of the new problem (17)-(23):

$$\overline{f(\boldsymbol{\gamma})} \geq f(\overline{\boldsymbol{\gamma}}) = f(\overline{\boldsymbol{x}})$$
$$\overline{g_l(\boldsymbol{\gamma})} \geq g_l(\overline{\boldsymbol{\gamma}}) = g_l(\overline{\boldsymbol{x}})$$

From this it easily follows that the minimum cost metric of the new problem (17)-(23) is also *greater than or equal* to that of the original problem (3)-(7).

Such auxiliary variables are introduced in [2] and [1] to optimize functions of time averages, which is very different from optimizing time averages of functions. Note that if $f(\cdot) = g_l(\cdot) = 0$ for all l, then $M = 0$ and we do not need any auxiliary variables.[2] Following the framework of [2] [1], which solves problems similar to the above under ergodic assumptions on $\omega(t)$, in addition to the actual queues $Q_k(t)$ we define *virtual queues* $Z_l(t)$ and $H_m(t)$ for each $l \in \{1, \ldots, L\}$ and $m \in \{1, \ldots, M\}$, with $Z_l(0) = H_m(0) = 0$ for all l and m, and with update equation:

$$Z_l(t + 1) = \max[Z_l(t) + y_l(t) + g_l(\boldsymbol{\gamma}(t)), 0] \qquad (24)$$

$$H_m(t + 1) = H_m(t) + \gamma_m(t) - x_m(t) \qquad (25)$$

Note that the queues $Q_k(t)$ and $Z_l(t)$ are non-negative for all t, while the queues $H_m(t)$ can be possibly negative. If the queues $Z_l(t)$, $Q_k(t)$, $H_m(t)$ are close to zero at time t_{end}, then the inequality constraints (18), (19), (20), respectively, are close to being satisfied, as specified by the following lemma.

Lemma 1: (Approximate Constraint Satisfaction) For any sequence $\{\omega(0), \omega(1), \omega(2), \ldots\}$, any designated end time

$t_{end} > 0$, any non-negative initial queue states $Z_l(0)$, $Q_k(0)$, any real valued initial queue states $H_m(0)$, and any sequence of control decisions $\alpha(0) \in \mathcal{A}_{\omega(0)}$, $\alpha(1) \in \mathcal{A}_{\omega(1)}$, $\alpha(2) \in \mathcal{A}_{\omega(2)}$, etc., we have for all $l \in \{1, \ldots, L\}$, $k \in \{1, \ldots, K\}$:

$$\overline{a}_k \ \leq \ \overline{b}_k + \frac{Q_k(t_{end}) - Q_k(0)}{t_{end}} \qquad (26)$$

$$\overline{y}_l + g_l(\overline{\boldsymbol{x}}) \ \leq \ \frac{Z_l(t_{end}) - Z_l(0)}{t_{end}}$$
$$+ \sum_{m=1}^{M} \frac{\beta_{l,m}|H_m(t_{end}) - H_m(0)|}{t_{end}} \qquad (27)$$

$$\overline{\boldsymbol{x}} + \boldsymbol{\epsilon} \in \mathcal{X} \qquad (28)$$

where $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_M)$ is a vector with components that satisfy:

$$|\epsilon_m| = \frac{|H_m(t_{end}) - H_m(0)|}{t_{end}} \quad \forall m \in \{1, \ldots, M\}$$

where $\overline{\boldsymbol{x}} = (\overline{x}_1, \ldots, \overline{x}_M)$ represents a time average over the first t_{end} slots, as defined in (2), as do time averages $\overline{a}_k$, $\overline{b}_k$, $\overline{y}_l$.

Note that the inequalities (26)-(28) correspond to the desired constraints (4)-(6), and show that these desired constraints are approximately satisfied if the values of $Q_k(t_{end})/t_{end}$, $Z_l(t_{end})/t_{end}$, $|H_m(t_{end})|/t_{end}$ are small. Recall that $\beta_{l,m}$ are bounds on the partial derivatives of $g_l(\boldsymbol{x})$, and hence if for a particular l we have $g_l(\boldsymbol{x}) = 0$ for all $\boldsymbol{x}$, then $\beta_{l,m} = 0$ for all m, which tightens the bound in (27). This is useful for linear cost functions, as described in more detail in Section III-D.

Proof: (Lemma 1) From the queueing update equation (1) we have for any $k \in \{1, \ldots, K\}$ and all $t \geq 0$:

$$Q_k(t + 1) \geq Q_k(t) - b_k(t) + a_k(t)$$

Summing over $\tau \in \{0, 1, \ldots, t - 1\}$ for a given $t > 0$, and dividing by t gives:

$$\frac{1}{t} \sum_{\tau=0}^{t-1} a_k(\tau) \leq \frac{1}{t} \sum_{\tau=0}^{t-1} b_k(\tau) + \frac{Q_k(t) - Q_k(0)}{t}$$

Plugging $t = t_{end}$ into the above inequality proves (26).

Similarly, the update equation (25) easily leads to the following for all $m \in \{1, \ldots, M\}$:

$$\left| \frac{1}{t} \sum_{\tau=0}^{t-1} \gamma_m(\tau) - \frac{1}{t} \sum_{\tau=0}^{t-1} x_m(\tau) \right| = \frac{|H_m(t) - H_m(0)|}{t}$$

Using $t = t_{end}$ in the above yields:

$$|\overline{\gamma}_m - \overline{x}_m| = \frac{|H_m(t_{end}) - H_m(0)|}{t_{end}} \qquad (29)$$

Note that $\boldsymbol{\gamma}(\tau) \in \mathcal{X}$ for all $\tau \in \{0, \ldots, t_{end} - 1\}$, and hence (by convexity of $\mathcal{X}$) the time average $\overline{\boldsymbol{\gamma}}$ is also in $\mathcal{X}$. Defining $\boldsymbol{\epsilon} \triangleq \overline{\boldsymbol{\gamma}} - \overline{\boldsymbol{x}}$ thus ensures that $\overline{\boldsymbol{x}} + \boldsymbol{\epsilon} \in \mathcal{X}$. Noting that $\epsilon_m = \overline{\gamma}_m - \overline{x}_m$ and using (29) proves (28).

Finally, from (24) we have for any $l \in \{1, \ldots, L\}$ and for all slots $t \geq 0$:

$$Z_l(t + 1) \geq Z_l(t) + y_l(t) + g_l(\boldsymbol{\gamma}(t))$$

---

[2]Also, in this case $M = 0$, we do not need any queues of the type (25).

Summing over $\tau \in \{0, \ldots, t-1\}$ for any time $t > 0$ yields:

$$Z_l(t) - Z_l(0) \geq \sum_{\tau=0}^{t-1} y_l(\tau) + \sum_{\tau=0}^{t-1} g_l(\boldsymbol{\gamma}(\tau))$$

Dividing by t and rearranging terms proves that:

$$\frac{1}{t} \sum_{\tau=0}^{t-1} [y_l(\tau) + g_l(\boldsymbol{\gamma}(\tau))] \leq \frac{Z_l(t) - Z_l(0)}{t} \qquad (30)$$

Using $t = t_{end}$ in the above inequality, together with Jensen's inequality for the convex function $g_l(\cdot)$, yields:

$$\overline{y}_l + g_l(\overline{\boldsymbol{\gamma}}) \leq \frac{Z_l(t_{end}) - Z_l(0)}{t_{end}} \qquad (31)$$

where $\overline{\boldsymbol{\gamma}}$ is the time average of $\boldsymbol{\gamma}(\tau)$ over the first t_{end} slots. However, by (13) we have:

$$g_l(\overline{\boldsymbol{x}}) \leq g_l(\overline{\boldsymbol{\gamma}}) + \sum_{m=1}^{M} \beta_{l,m} |\overline{x}_m - \overline{\gamma}_m|$$

Using this and (29) in (31) proves (27). $\qquad \square$

A. Quadratic Lyapunov Functions and Sample Path Drift

Lemma 1 ensures the constraints (18)-(23) (and hence the constraints (4)-(7)) are approximately satisfied if the final queue states of all queues are small relative to t_{end}. Define $\boldsymbol{\Theta}(t)$ as a vector of all current queue states:

$$\boldsymbol{\Theta}(t) \triangleq [\boldsymbol{Z}(t), \boldsymbol{Q}(t), \boldsymbol{H}(t)]$$

where $\boldsymbol{Z}(t)$, $\boldsymbol{Q}(t)$, $\boldsymbol{H}(t)$ are vectors with entries $Z_l(t)$, $Q_k(t)$, $H_m(t)$, respectively. As a scalar measure of queue size, define the following quadratic Lyapunov function, as in [1]:

$$L(\boldsymbol{\Theta}(t)) \triangleq \frac{1}{2} \sum_{l=1}^{L} Z_l(t)^2 + \frac{1}{2} \sum_{k=1}^{K} Q_k(t)^2 + \frac{1}{2} \sum_{m=1}^{M} H_m(t)^2 \quad (32)$$

For a given positive integer T, let $\Delta_T(t)$ represent the T-slot sample path Lyapunov drift associated with particular controls implemented over the interval $\{t, \ldots, t+T-1\}$ when the queues have state $\boldsymbol{\Theta}(t)$ at the beginning of the interval:[3]

$$\Delta_T(t) \triangleq L(\boldsymbol{\Theta}(t+T)) - L(\boldsymbol{\Theta}(t)) \qquad (33)$$

This notion of T-slot drift differs from that given in [1] in that it does not involve an expectation. It is difficult to control the T-slot drift, because it depends on future (and hence unknown) $\omega(t)$ values. Thus, following the approach in [1], we design a control policy that, every slot t, observes the current $\omega(t)$ value and the current queue states $\boldsymbol{\Theta}(t)$ and chooses a control action $\alpha(t) \in \mathcal{A}_{\omega(t)}$ to minimize a weighted sum of the 1-slot drift and the current contribution to the cost metric (17):

Minimize: $\quad \Delta_1(t) + V\hat{y}_0(\alpha(t), \omega(t)) + Vf(\boldsymbol{\gamma}(t))$

Subject to: $\qquad$ Constraints (21)-(23)

where $V \geq 0$ is a control parameter chosen in advance that affects a performance tradeoff. Rather than perform the exact

[3]Note that the value of $\Delta_T(t)$ depends on the queue state $\boldsymbol{\Theta}(t)$ at the start of the T-slot interval, the random events $\{\omega(t), \ldots, \omega(t+T-1)\}$, and the control actions $\{\alpha(t), \ldots, \alpha(t+T-1)\}$ that are chosen over this interval.

minimization of the above problem, it suffices to minimize a bound. The following lemma bounds $\Delta_1(t)$.

Lemma 2: Under Assumption A1, the 1-slot drift $\Delta_1(t)$ satisfies:

$$\Delta_1(t) \leq B + \sum_{l=1}^{L} Z_l(t)[y_l(t) + g_l(\boldsymbol{\gamma}(t))]$$
$$+ \sum_{k=1}^{K} Q_k(t)[a_k(t) - b_k(t)]$$
$$+ \sum_{m=1}^{M} H_m(t)[\gamma_m(t) - x_m(t)]$$

where B is a finite constant that satisfies for all t:

$$B \geq \frac{1}{2} \sum_{l=1}^{L} (y_l(t) + g_l(\boldsymbol{\gamma}(t)))^2 + \frac{1}{2} \sum_{k=1}^{K} [b_k(t)^2 + a_k(t)^2]$$
$$+ \frac{1}{2} \sum_{m=1}^{M} (\gamma_m(t) - x_m(t))^2 \qquad (34)$$

Such a finite constant B exists by the boundedness assumptions A1, and a particular such B is given in Appendix E.

Proof: Squaring the $Z_l(t)$ update equation (24) and noting that $\max[x, 0]^2 \leq x^2$ yields:

$$Z_l(t+1)^2 \leq Z_l(t)^2 + (y_l(t) + g_l(\boldsymbol{\gamma}(t)))^2$$
$$+ 2Z_l(t)[y_l(t) + g_l(\boldsymbol{\gamma}(t))] \qquad (35)$$

Similarly, from (1) we have:

$$Q_k(t+1)^2 \leq (Q_k(t) - b_k(t))^2 + a_k(t)^2$$
$$+ 2a_k(t)\max[Q_k(t) - b_k(t), 0]$$
$$\leq (Q_k(t) - b_k(t))^2 + a_k(t)^2 + 2a_k(t)Q_k(t)$$
$$= Q_k(t)^2 + a_k(t)^2 + b_k(t)^2$$
$$+ 2Q_k(t)[a_k(t) - b_k(t)] \qquad (36)$$

Finally, from (25) we have:

$$H_m(t+1)^2 = H_m(t)^2 + (\gamma_m(t) - x_m(t))^2$$
$$+ 2H_m(t)(\gamma_m(t) - x_m(t)) \qquad (37)$$

Summing (35), (36), (37) and dividing by 2 yields the result. $\qquad \square$

The above lemma shows that:

$$\Delta_1(t) + V\hat{y}_0(\alpha(t), \omega(t)) + Vf(\boldsymbol{\gamma}(t)) \leq$$
$$B + V\hat{y}_0(\alpha(t), \omega(t)) + Vf(\boldsymbol{\gamma}(t))$$
$$+ \sum_{l=1}^{L} Z_l(t)[\hat{y}_l(\alpha(t), \omega(t)) + g_l(\boldsymbol{\gamma}(t))]$$
$$+ \sum_{k=1}^{K} Q_k(t)[\hat{a}_k(\alpha(t), \omega(t)) - \hat{b}_k(\alpha(t), \omega(t))]$$
$$+ \sum_{m=1}^{M} H_m(t)[\gamma_m(t) - \hat{x}_m(\alpha(t), \omega(t))] \qquad (38)$$

B. The General Universal Scheduling Algorithm

Our universal scheduling algorithm is designed to minimize the right hand side of (38) every slot, as described as follows: Every slot t, observe the random event $\omega(t)$ and the current queue backlogs $Z_l(t)$, $Q_k(t)$, $H_m(t)$ for all $l \in \{1, \ldots, L\}$, $k \in \{1, \ldots, K\}$, $m \in \{1, \ldots, M\}$, and perform the following.

- Choose $\boldsymbol{\gamma}(t) = (\gamma_1(t), \ldots, \gamma_M(t))$ to solve the following:

  Minimize: $\quad V f(\boldsymbol{\gamma}(t)) + \sum_{l=1}^{L} Z_l(t) g_l(\boldsymbol{\gamma}(t))$
  $$+ \sum_{m=1}^{M} H_m(t) \gamma_m(t)$$

  Subject to: $\quad\quad\quad \boldsymbol{\gamma}(t) \in \mathcal{X}$
  $$x_m^{min} \le \gamma_m(t) \le x_m^{max} \quad \forall m \in \{1, \ldots, M\}$$

- Choose $\alpha(t) \in \mathcal{A}_{\omega(t)}$ to minimize:

$$V \hat{y}_0(\alpha(t), \omega(t)) + \sum_{l=1}^{L} Z_l(t) \hat{y}_l(\alpha(t), \omega(t))$$
$$- \sum_{m=1}^{M} H_m(t) \hat{x}_m(\alpha(t), \omega(t))$$
$$+ \sum_{k=1}^{K} Q_k(t) [\hat{a}_k(\alpha(t), \omega(t)) - \hat{b}_k(\alpha(t), \omega(t))] \quad (39)$$

- Update the actual queues $Q_k(t)$ and the virtual queues $Z_l(t)$, $H_m(t)$ for all $k \in \{1, \ldots, K\}$, $l \in \{1, \ldots, L\}$, $m \in \{1, \ldots, M\}$ via (1), (24), and (25).

Note that the above selection of $\boldsymbol{\gamma}(t)$ minimizes a convex function over a convex set, and decomposes into M decoupled convex optimizations of one variable in the case when cost functions $f(\boldsymbol{x})$ and $g_l(\boldsymbol{x})$ have the separable structure of (14) and when the set $\mathcal{X}$ is equal to $\mathbb{R}^N$ or a hypercube in $\mathbb{R}^N$. The optimization of $\alpha(t)$ in (39) may be more complex and is possibly a non-convex or combinatorial problem (depending on the $\mathcal{A}_{\omega(t)}$ set and the $\hat{x}_m(\cdot)$, $\hat{y}_l(\cdot)$, $\hat{b}_k(\cdot)$, and $\hat{a}_k(\cdot)$ functions). However, it is simple when the action space $\mathcal{A}_{\omega(t)}$ contains only a finite (and small) number of options, in which case we can simply compare the functional (39) for each option. A key property of the above algorithm is that it is *non-anticipating* in that it acts only on the current $\omega(t)$ value, without knowledge of future values.

It can be shown that the expression (39) has a well defined minimum value over the set $\mathcal{A}_{\omega(t)}$ whenever Assumption A4 in Appendix B holds. However, the next theorem allows for *approximate* minimization, where the choice of decision variables $\boldsymbol{\gamma}(t)$ and $\alpha(t)$ lead to a value that is off by an additive constant from achieving the minimum (or infimum) of the right hand side in (38). This is similar to the approximation results in [1] and references therein, developed for ergodic problems. Specifically, we define an algorithm to be *C-approximate* if every slot it makes decisions $\boldsymbol{\gamma}(t)$ and $\alpha(t)$ to satisfy the constraints (21)-(23) and to yield either the infimum of the expression on the right hand side of (38) (as described in the algorithm above), or to yield a value on the right hand side that differs from the infimum by at most an additive constant $C \ge 0$.

Theorem 1: Suppose Assumptions A1 and A2 hold. Consider any C-approximate algorithm. Let the random event sequence $\{\omega(0), \omega(1), \omega(2), \ldots\}$ be arbitrary. Then:

(a) For any slot $t > 0$ we have:

$$\sum_{l=1}^{L} Z_l(t)^2 + \sum_{k=1}^{K} Q_k(t)^2 + \sum_{m=1}^{M} H_m(t)^2 \le tVC_0^2 + 2L(\boldsymbol{\Theta}(0))$$

where the constant C_0 is defined:

$$C_0 \triangleq \sqrt{2[(B+C)/V + (y_0^{max} - y_0^{min}) + (f^{max} - f^{min})]}$$

where B is the non-negative constant defined in (34). In particular, all queues are bounded as follows:

$$Z_l(t), Q_k(t), |H_m(t)| \le \sqrt{tVC_0^2 + 2L(\boldsymbol{\Theta}(0))}$$

This bound becomes $C_0\sqrt{tV}$ if all queues are initially empty.

(b) If all queues are initially empty, then for any designated time $t_{end} > 0$ we have:

$$\overline{y}_l + g_l(\overline{\boldsymbol{x}}) \le C_0 \sqrt{\frac{V}{t_{end}}} \left[1 + \sum_{m=1}^{M} \beta_{l,m} \right] \quad \forall l \in \{1, \ldots, L\}$$
$$\overline{a}_k \le \overline{b}_k + C_0 \sqrt{\frac{V}{t_{end}}} \quad\quad\quad \forall k \in \{1, \ldots, K\}$$
$$\overline{\boldsymbol{x}} + \boldsymbol{\epsilon}(t) \in \mathcal{X}$$

where $\boldsymbol{\epsilon}(t) = (\epsilon_1(t), \ldots, \epsilon_M(t))$ has entries that satisfy:

$$|\epsilon_m(t)| \le C_0 \sqrt{\frac{V}{t_{end}}}$$

Thus, the constraints (7) of the original problem are satisfied, and the constraints (4)-(6) of the original problem are approximately satisfied, where the error term in the approximation decays with t_{end} according to a constant multiple of $\sqrt{V/t_{end}}$.

(c) Consider any positive integer frame size T, any positive integer R, and define $t_{end} = RT$. Then the value of the system cost metric over t_{end} slots satisfies:

$$\overline{y}_0 + f(\overline{\boldsymbol{x}}) \le \frac{1}{R} \sum_{r=0}^{R-1} F_r^* + \frac{B+C}{V} + \frac{D(T-1)}{V}$$
$$+ \frac{L(\boldsymbol{\Theta}(0))}{VRT} + \sum_{m=1}^{M} \frac{\nu_m |H_m(RT) - H_m(0)|}{RT} \quad (40)$$

where $\overline{y}_m$, $\overline{\boldsymbol{x}}$ are time averages over the first t_{end} slots, and where F_r^* is the optimal solution to the problem (15) and represents the optimal cost achieved by an idealized T-slot lookahead policy implemented over the rth frame of size T. The constant D is a finite constant that satisfies for all t and all possible control actions that can be implemented on slot t:

$$D \ge \frac{1}{2} \sum_{l=1}^{L} z_l^{diff} |y_l(t) + g_l(\boldsymbol{\gamma}(t))|$$
$$+ \sum_{m=1}^{M} h_m^{diff} |x_m(t) - \gamma_m(t)|$$
$$+ \frac{1}{2} \sum_{k=1}^{K} q_k^{diff} \max[b_k(t), a_k(t)] \quad (41)$$

where z_l^{diff}, q_k^{diff}, h_m^{diff} represent the maximum change in queues $Z_l(t)$, $Q_k(t)$, $H_m(t)$ over one slot, given by:

$$z_l^{diff} \triangleq \max[|y_l^{max} + g_l^{max}|, |y_{l,min} + g_l^{min}|]$$
$$q_k^{diff} \triangleq \max[b_k^{max}, a_k^{max}]$$
$$h_m^{diff} \triangleq |x_m^{max} - x_m^{min}|$$

Such a finite constant exists by the boundedness assumptions, and a value of D that satisfies inequality (41) is given in Appendix E.

Finally, if initial queue backlogs are 0, the final term in (40) is bounded by:

$$\sum_{m=1}^{M} \frac{\nu_m |H_m(RT) - H_m(0)|}{RT} \leq \sum_{m=1}^{M} \nu_m C_0 \sqrt{\frac{V}{RT}}$$

Proof: See Appendix A. $\qquad \square$

C. Discussion of Theorem 1

Consider the simple case when queues are initially 0, $g_l(\boldsymbol{x}) = 0$ for all l (so that $\beta_{l,m} = 0$), and where we use a C-approximate algorithm for some constant $C \geq 0$. Fix a positive integer frame size T. Theorem 1 can be interpreted as follows: The algorithm implemented over $t_{end} = RT$ slots ensures that the desired constraints (4)-(6) are approximately met to within a "fudge factor" given by:

$$\text{Constraint fudge factor} = C_0 \sqrt{\frac{V}{RT}}$$

where C_0 is the constant defined in part (a) of Theorem 1. This fudge factor is made arbitrarily small when RT is sufficiently larger than V. Further, the achieved cost is either smaller than the cost $\frac{1}{R} \sum_{r=0}^{R-1} F_r^*$ associated with an ideal T-slot lookahead policy implemented over R successive frames of length T, or differs from this value by an amount no more than a fudge factor that satisfies:

$$\text{Cost fudge factor} = \frac{C_1 T}{V} + C_2 \sqrt{\frac{V}{RT}}$$

The value of V can be chosen so that $C_1 T/V$ is arbitrarily small, in which case both the cost fudge factor and the constraint fudge factor are arbitrarily small provided that R is sufficiently large, that is, provided that we wait for a sufficiently large number of frames.

The constants C_1 and C_2 are given by:

$$
\begin{aligned}
C_1 &\triangleq (B + C - D)/T + D \\
C_2 &\triangleq C_0 \sum_{m=1}^{M} \nu_m
\end{aligned}
$$

Note that $C_2 = 0$ in the case when $f(\boldsymbol{x}) = 0$ (so that $\nu_m = 0$ for all m). Finally, note that the value of T does not need to be chosen in order to implement the algorithm (we need only choose a value of V), and hence the above bounds can be optimized over all positive integers T.

The above tradeoffs described by V and R hold for general problems of the type (3)-(7), and can be tightened for particular problems such as the network problem described in the next section, which provides queue bounds that do not grow with time. A similar strengthening due to constant queue bounds can be shown for the general problem in the case when Assumption A1 is strengthened to a "Slater-type" condition, as described in Section VI.

D. Linear Cost Functions

The auxiliary variables are crucial for optimization of time varying systems with non-linear cost functions $f(\boldsymbol{x})$, $g_l(\boldsymbol{x})$. However, they are not needed when cost functions are linear (or affine). For example, let $\boldsymbol{x}(t)$ be a vector of attributes as defined before, and consider the problem:

$$
\begin{aligned}
\text{Minimize:} \quad & h_0(\overline{\boldsymbol{x}}) \\
\text{Subject to:} \quad & h_l(\overline{\boldsymbol{x}}) \leq 0 \quad \forall l \in \{1, \ldots, L\} \\
& \alpha(t) \in \mathcal{A}_{\omega(t)} \quad \forall t \in \{0, \ldots, t_{end} - 1\}
\end{aligned}
$$

where $h_l(\boldsymbol{x})$ are affine functions (i.e., linear plus a constant), so that $\overline{h_l(\boldsymbol{x})} = h_l(\overline{\boldsymbol{x}})$. This can of course be treated using the framework of (3)-(7) with $h_0(\boldsymbol{x}) = f(\boldsymbol{x})$, and $g_l(\boldsymbol{x}) = h_l(\boldsymbol{x})$, $y_l(\boldsymbol{x}) = 0$. However, it can also be treated using (3)-(7) with $f(\boldsymbol{x}) = g_l(\boldsymbol{x}) = 0$ for all $l \in \{1, \ldots, L\}$ and $y_l(\boldsymbol{x}) = h_l(\boldsymbol{x})$ (noting that $\overline{y}_l = h_l(\overline{\boldsymbol{x}})$). This latter method is advantageous because it has $\nu_m = \beta_{l,m} = 0$ for all l and m, which tightens the constraint inequalities (27) and the cost guarantee (40). Thus, it is useful to exploit linearity whenever possible.

E. Infinite Horizons and Ergodicity

Consider now the problem (3)-(7) over an *infinite horizon*, so that time averages $\overline{y}_l$, $\overline{\boldsymbol{x}}$, $\overline{a}_k$, $\overline{b}_k$ represent limiting averages over the infinite horizon. Suppose that Assumptions A1 and A2 hold, and that we use a C-approximate algorithm so that Theorem 1 applies. By taking a limit as $t_{end} \to \infty$, Theorem 1b implies that all required infinite horizon constraints are met. Specifically, we have:

$$\limsup_{t \to \infty} \left[\overline{y}_l(t) + g_l(\overline{\boldsymbol{x}}(t)) \right] \leq 0 \quad \forall l \in \{1, \ldots, L\} \quad (42)$$

$$\limsup_{t \to \infty} \left[\overline{a}_k(t) - \overline{b}_k(t) \right] \leq 0 \quad \forall k \in \{1, \ldots, K\} \quad (43)$$

$$\lim_{t_i \to \infty} \overline{\boldsymbol{x}}(t_i) \in \mathcal{X} \quad (44)$$

where $\overline{y}_l(t)$, $\overline{\boldsymbol{x}}(t)$, $\overline{a}_k(t)$, $\overline{b}_k(t)$ represent time averages over the first t slots, and where $\{t_i\}$ is any subsequence of times over which $\overline{\boldsymbol{x}}(t)$ converges.[4] Further, Theorem 1c implies that the infinite horizon cost satisifes:

$$
\begin{aligned}
\lim_{R \to \infty} \left[\overline{y}_0(RT) + f(\overline{\boldsymbol{x}}(RT)) \right] \leq{} & \lim_{R \to \infty} \frac{1}{R} \sum_{r=0}^{R-1} F_r^* \\
& + \frac{B+C}{V} + \frac{D(T-1)}{V} \quad (45)
\end{aligned}
$$

Consider now the special case when the random events $\{\omega(0), \omega(1), \omega(2), \ldots\}$ evolve according to a general ergodic process with a well defined time average probability distribution. In this case and under some mild assumptions, it can be shown that the optimal infinite horizon cost f^* can be achieved over the class of stationary and randomized algorithms, that F_r^* is close to f^* for each r whenever T is sufficiently large, and that the term $\frac{1}{R} \sum_{r=0}^{R-1} F_r^*$ converges to f^* plus an error term that is bounded by $\delta(T)$, where $\delta(T)$ is a function that satisfies $\lim_{T \to \infty} \delta(T) = 0$. This is discussed in more detail in Appendix C.

---

[4] Note that $\overline{\boldsymbol{x}}(t)$ is an infinite sequence (with time index t) that takes values in a compact set, and so it has a convergent subsequence.

IV. A Simple Internet Model

Here we apply the universal scheduling framework to a simple flow based internet model, where we neglect the actual network queueing and develop a flow control policy that simply ensures the flow rate over any link is not more than the link capacity (similar to the flow based models in [23][24][25][26][27]). Section V treats a more extensive network model that explicitly accounts for all queues.

Suppose there are N nodes and L links, where each link $l \in \{1, \ldots, L\}$ has a possibly time-varying link capacity $C_l(t)$, for slotted time $t \in \{0, 1, 2, \ldots\}$. Suppose there are M sessions, and let $A_m(t)$ represent the new arrivals to session m on slot t. Each session $m \in \{1, \ldots, M\}$ has a particular source node and a particular destination node. Assume link capacities and newly arriving traffic are bounded so that:

$$0 \leq C_l(t) \leq C_l^{max} \;\; \forall t \;\;,\;\; 0 \leq A_m(t) \leq A_m^{max} \;\; \forall t \quad (46)$$

for some finite constants C_l^{max} and A_m^{max}. The random network event $\omega(t)$ is thus given by:

$$\omega(t) \triangleq [(C_1(t), \ldots, C_L(t)); (A_1(t), \ldots, A_M(t))] \quad (47)$$

Recall that $\omega(t)$ is an arbitrary sequence with no probability model. The control action taken every slot is to first choose $x_m(t)$, the amount of type m traffic admitted into the network on slot t, according to:

$$0 \leq x_m(t) \leq A_m(t)$$

Next, we must specify a *path* for this data, from a collection of paths $\mathcal{P}_m(t)$ associated with path options of session m on slot t (possibly being the set of all possible paths in the network from the source of session m to its destination).[5] Here, a path is defined in the usual sense, being a sequence of links starting at the source, ending at the destination, and being such that the end node of each link is the start node of the next link. Let $1_{l,m}(t)$ be an indicator variable that is 1 if the data $x_m(t)$ is selected to use a path that contains link l. The $(1_{l,m}(t))$ values completely specify the chosen paths for slot t, and hence the decision variable for slot t is given by:

$$\alpha(t) \triangleq [(x_1(t), \ldots, x_M(t)); (1_{l,m}(t))|_{l \in \{1, \ldots, L\}, m \in \{1, \ldots, M\}}]$$

Let $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_M)$ be a vector of the infinite horizon time average admitted flow rates, and let $\phi(x) = \sum_{m=1}^{M} \phi_m(x_m)$ be a separable *utility function*. Assume that each $\phi_m(x)$ is a continuous, concave, non-decreasing function in x, with maximum right derivative $\nu_m < \infty$. Our goal is to maximize the throughput-utility $\phi(\overline{x})$ subject to the constraints that the time average flow over each link l is less than or equal to the time average capacity of that link. The infinite horizon utility optimization problem of interest is thus:

Maximize: $\quad \sum_{m=1}^{M} \phi_m(\overline{x}_m)$

Subject to: $\quad \sum_{m=1}^{M} \overline{1_{l,m}x_m} \leq \overline{C}_l \;\; \forall l \in \{1, \ldots, L\}$

[5]Strictly speaking, if there are time varying path choices then we should augment $\omega(t)$ in (47) to include $\mathcal{P}_m(t)$ for all $m \in \{1, \ldots, M\}$.

where the time averages are defined:

$$\overline{x}_m \;\; \triangleq \;\; \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} x_m(\tau)$$

$$\overline{1_{l,m}x_m} \;\; \triangleq \;\; \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} 1_{l,m}(\tau) x_m(\tau)$$

$$\overline{C}_l \;\; \triangleq \;\; \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} C_l(\tau)$$

This is equivalent to minimizing the convex function $f(\overline{x}) = -\phi(\overline{x})$, and hence exactly fits our framework. There is no set constraint (6), so that $\mathcal{X} = \mathbb{R}^M$. As there are no actual queues $Q_k(t)$ in this model, we use only virtual queues $Z_l(t)$ and $H_m(t)$, defined by update equations:

$$Z_l(t+1) = \max[Z_l(t) + \sum_{m=1}^{M} 1_{l,m}(t) x_m(t) - C_l(t), 0] \quad (48)$$

$$H_m(t+1) = H_m(t) + \gamma_m(t) - x_m(t) \quad (49)$$

where $\gamma_m(t)$ are auxiliary variables for $m \in \{1, \ldots, M\}$. This is equivalent to the general framework with $y_l(t) = \sum_{m=1}^{M} 1_{l,m}(t) x_m(t) - C_l(t)$ and $g_l(\cdot) = 0$ for $l \in \{1, \ldots, L\}$. Note that Assumption A1 holds by the boundedness assumptions (46), and Assumption A2 holds because this system has an "idle" control action that admits no new data (so that $y_l(t) \leq 0$ for all l under this idle action). The general universal scheduling algorithm for this problem thus reduces to:

- (Auxiliary Variables) Every slot t, each session $m \in \{1, \ldots, M\}$ observes $H_m(t)$ and chooses $\gamma_m(t)$ as the solution to:

  Maximize: $\quad V\phi_m(\gamma_m(t)) - H_m(t)\gamma_m(t) \quad (50)$
  Subject to: $\quad 0 \leq \gamma_m(t) \leq A_m^{max} \quad (51)$

  This is a simple maximization of a concave single variable function over an interval.

- (Routing and Flow Control) For each slot t and each session $m \in \{1, \ldots, M\}$, observe the new arrivals $A_m(t)$, the queue backlog $H_m(t)$, and the link queues $Z_l(t)$, and choose $x_m(t)$ and a path to maximize:

  Maximize: $\quad x_m(t)H_m(t) - x_m(t)\sum_{l=1}^{L} 1_{l,m}(t)Z_l(t)$
  Subject to: $\quad 0 \leq x_m(t) \leq A_m(t)$
  $\quad$ The path specified by $(1_{l,m}(t))$ is in $\mathcal{P}_m(t)$

  This reduces to the following: First find a shortest path from the source of session m to the destination of session m, using link weights $Z_l(t)$ as link costs. If the total weight of the shortest path is less than or equal to $H_m(t)$, then choose $x_m(t) = A_m(t)$ and route all of this data over this single shortest path. Else, there is too much congestion in the network, and so we choose $x_m(t) = 0$ (thereby dropping all data $A_m(t)$).

- (Virtual Queue Updates) Update the virtual queues according to (48) and (49).

The shortest path routing in this algorithm is similar to that given in [26], which treats a flow-based network stability

problem in an ergodic setting under the assumption that arriving traffic is admissible (so that flow control is not used).

Assume for simplicity that all queues are initially empty. Then for any frame size T and any number of frames R, from Theorem 1c, we know the utility of this algorithm satisfies:

$$\phi(\overline{x}) \geq \frac{1}{R} \sum_{r=1}^{R-1} \Phi_r^* - \frac{B}{V} - \frac{D(T-1)}{V}$$
$$- \sum_{m=1}^{M} \frac{\nu_m |H_m(RT)|}{RT} \quad (52)$$

where $\overline{x}$ represents a time average over the first RT slots, and Φ_r^* represents the utility achieved by the T-slot lookahead policy implemented over slots $\{rT, \ldots, rT+T-1\}$. Here we are assuming we use an exact implementation of the algorithm (a 0-approximation), so that $C = 0$. The constants B and D given in (34) and (41) are simplified for this context without queues $Q_k(t)$, and are provided in Appendix E.

Furthermore, the infinite horizon constraints are satisfied by (42), and bounds on the virtual queue sizes for any time $t > 0$ are also given in Theorem 1. However, with this particular structure we can obtain tighter bounds. Indeed, from the update equation for $H_m(t)$ in (49) and the auxiliary variable algorithm defined in (50)-(51), it is easy to see that:

- If $H_m(t) < 0$, then $\gamma_m(t) = A_m^{max}$ and hence $H_m(t)$ cannot decrease on the next slot.
- If $H_m(t) > V\nu_m$, then $\gamma_m(t) = 0$ and hence $H_m(t)$ cannot increase on the next slot.

It easily follows that:

$$-A_m^{max} \leq H_m(t) \leq V\nu_m + A_m^{max} \quad \forall t \in \{0, 1, 2, \ldots\} \quad (53)$$

provided that this is true for $H_m(0)$ (which is indeed the case if $H_m(0) = 0$). Therefore, the final term in the utility guarantee (52) is bounded by:

$$\sum_{m=1}^{M} \frac{\nu_m |H_m(RT)|}{RT} \leq \sum_{m=1}^{M} \frac{\nu_m(V\nu_m + A_m^{max})}{RT}$$

which goes to zero as the number of frames R goes to infinity. We further note that the utility guarantee (52) can be modified to apply to *any interval of RT slots, starting at any slot t_0*, provided that we modify the equation to account for possibly non-zero initial queue conditions according to (40).

Further, the fact that the queues $H_m(t)$ are deterministically bounded allows one to deterministically bound the queue sizes $Z_l(t)$ as follows: For all $l \in \{1, \ldots, L\}$ we have:

$$0 \leq Z_l(t) \leq V\nu^{max} + (M+1)A^{max} \quad \forall t \quad (54)$$

where ν^{max} and A^{max} are defined as the maximum of all ν_m and A_m^{max} values:

$$\nu^{max} \triangleq \max_{m \in \{1, \ldots, M\}} \nu_m \quad , \quad A^{max} \triangleq \max_{m \in \{1, \ldots, M\}} A_m^{max}$$

The proof of this fact is simple: If a link l satisfies $Z_l(t) \leq V\nu^{max} + A^{max}$, then on the next slot we have $Z_l(t+1) \leq V\nu^{max} + (M+1)A^{max}$, because the queue can increase by at most MA^{max} on any slot (see update equation (48)). Else, if $Z_l(t) > V\nu^{max} + A^{max}$, then any path that uses this link

incurs a cost larger than $V\nu^{max} + A^{max}$, which is larger than $H_m(t)$ for any session m. Thus, by the routing and flow control algorithm, no session will choose a path that uses this link on the current slot, and so $Z_l(t)$ cannot increase on the next slot.

A. Delayed Feedback

We note that it may be difficult to use the exact queue values $Z_l(t)$ when solving for the shortest path, as these values change every slot. Hence, a practical implementation may use out-of-date values $Z_l(t-\tau_{l,t})$ for some time delay $\tau_{l,t}$ that may depend on l and t. Further, the virtual queue updates for $Z_l(t)$ in (48) are most easily done at each link l, in which case the actual admitted data $x_m(t)$ for that link may not be known until some time delay, arriving as a process $x_m(t - \tau_{l,m,t})$. However, as the virtual queue size cannot change by more than a fixed amount every slot, the queue value used differs from the ideal queue value by no more than an additive constant that is proportional to the maximum time delay. In this case, provided that the maximum time delay is bounded, we are simply using a C-approximation and the utility and queue bounds are adjusted accordingly. A more extensive treatment of delayed feedback for the case of networks without dynamic arrivals or channels is found in [27], which uses a differential equation method.

B. Treating Wireless Networks with this Model

The above model can be applied equally to wireless networks. However, an important extension in this case is to allow the link capacities $(C_1(t), \ldots, C_L(t))$ to be functions of a network resource allocation decision (this is treated more extensively in Section V). This resource allocation can be viewed as part of the network control action taken every slot. It is not difficult to show from the general solution in Section III-B that the optimal resource allocation decision should observe $Z_l(t)$ values and choose capacities $C_l(t)$ to maximize the following weighted sum:

$$\sum_{l=1}^{L} C_l(t)Z_l(t)$$

Depending on the network model, this maximization can be difficult, and is generally prohibitively complex for wireless networks with interference. Fortunately, it is easy to show that if the attempted max-weight solution comes within a factor θ of the optimum max-weight decision every slot (for some value θ such that $0 < \theta \leq 1$), then the same utility guarantees hold with Φ_r^* re-defined as the optimal T-slot lookahead utility in a network with link capacities that are reduced by a factor θ from their actual values. This follows easily by noting that such a θ-*multiplicative-approximate* algorithm yields a right-hand-side in the drift bound (38) that is less than or equal to the right-hand-side associated with the optimal max-weight decisions implemented on a network with θ-reduced capacities (see also [1] for a more detailed discussion of this for the case of i.i.d. $\omega(t)$ events).

A. Transmission Variables

Let $S(t)$ represent the *topology state* of the network on slot t, observed on each slot t as in [1]. The value of $S(t)$ is an abstract and possibly multi-dimensional quantity that describes the current link conditions between all nodes under the current slot. The collection of all transmission rates that can be offered over each link (i, j) of the network is given by a general transmission rate function $\boldsymbol{C}(I(t), S(t))$:[7]

$$\boldsymbol{C}(I(t), S(t)) = (C_{ij}(I(t), S(t)))_{i,j \in \{1,\dots,N\}, i \neq j}$$

where $I(t)$ is a general network-wide resource allocation decision (such as link scheduling, bandwidth selection, modulation, etc.) and takes values in some abstract set $\mathcal{I}_{S(t)}$ that possibly depends on the current $S(t)$. We assume that the transmission rate function $C_{ij}(I(t), S(t))$ is non-negative and bounded by a finite constant μ_{ij}^{max} for all (i, j), $I(t)$, and $S(t)$.

Every slot the network controller observes the current $S(t)$ and makes a resource allocation decision $I(t) \in \mathcal{I}_{S(t)}$. The controller then chooses $\mu_{ij}^{(c)}(t)$ variables subject to the following constraints:

$$\mu_{ij}^{(c)}(t) \geq 0 \quad \forall i, j, c \in \{1, \dots, N\} \quad (57)$$

$$\mu_{ii}^{(c)}(t) = \mu_{ij}^{(i)}(t) = 0 \quad \forall i, j, c \in \{1, \dots, N\} \quad (58)$$

$$\sum_{c=1}^{N} \mu_{ij}^{(c)} \leq C_{ij}(I(t), S(t)) \quad \forall i, j \in \{1, \dots, N\} \quad (59)$$

B. The Utility Optimization Problem

This problem fits the general model of Section II by defining the random event $\omega(t)$ as follows:

$$\omega(t) \triangleq [\boldsymbol{A}(t); S(t)]$$

That is, the random event $\omega(t)$ is the collection of all new arrivals together with the current topology state. The control action $\alpha(t)$ is defined by:

$$\alpha(t) \triangleq [I(t); (\mu_{ij}^{(c)}(t))|_{i,j,c \in \{1,\dots,N\}}; (x_m(t))|_{m=1}^{M}]$$

representing the resource allocation, transmission, and flow control decisions. The action space $\mathcal{A}_{\omega(t)}$ is defined by the set of all $I(t) \in \mathcal{I}_{S(t)}$, all $(\mu_{ij}^{(c)}(t))$ that satisfy (57)-(59), and all $(x_m(t))$ that satisfy (55).

Define $\overline{x}_m$ as the time average of $x_m(t)$ over the first t_{end} slots (as in (2)), and define $\overline{\boldsymbol{x}}$ as the vector of these time averages. Our objective is to solve the following problem:

Maximize: $\quad \phi(\overline{\boldsymbol{x}}) \quad (60)$

Subject to: $\quad \alpha(t) \in \mathcal{A}_{\omega(t)} \ \forall t \in \{0, \dots, t_{end} - 1\} \quad (61)$

$$\sum_{m \in \mathcal{M}_n^{(c)}} \overline{x}_m + \sum_{i=1}^{N} \overline{\mu}_{in}^{(c)} \leq \sum_{j=1}^{N} \overline{\mu}_{nj}^{(c)}$$
$$\forall n, c \in \{1, \dots, N\} \quad (62)$$

[7]It is worth noting now that for networks with orthogonal channels, our "max-weight" transmission algorithm (to be defined in the next subsection) decouples to allow nodes to make transmission decisions based only on those components of the current topology state $S(t)$ that relate to their own local channels. Of course, for wireless interference networks, all channels are coupled, although distributed approximations of max-weight transmission exist in this case [1].

where $\phi(\overline{\boldsymbol{x}})$ is a continuous, concave, and entrywise non-decreasing utility function of the form:

$$\phi(\boldsymbol{x}) \triangleq \sum_{m=1}^{M} \phi_m(x_m)$$

Define ν_m as the right partial derivative of $\phi_m(x)$ at $x = 0$, and assume $0 \leq \nu_m < \infty$ for all m. Thus, this problem fits exactly into the general framework, satisfying Assumption A1 by our boundedness assumptions, and satisfying Assumption A2 by the "idle" control action that admits no new data and transmits over no links.

C. The Universal Network Scheduling Algorithm

To apply the general solution, note that the constraints (62) are upheld by stabilizing the actual queues $Q_n^{(c)}(t)$ with updates (56). Because we have not specified any additional constraints, there are no $Z_l(t)$ queues as used in the general framework. However, we have auxiliary variables $\gamma_m(t)$ for each $m \in \{1, \dots, M\}$, with virtual queue update:

$$H_m(t + 1) = H_m(t) + \gamma_m(t) - x_m(t) \quad (63)$$

The algorithm is thus:

- (Auxiliary Variables) For each slot t, each session $m \in \{1, \dots, M\}$ observes the current virtual queue $H_m(t)$, and chooses auxiliary variable $\gamma_m(t)$ as the solution to:

  Maximize: $\quad V\phi_m(\gamma_m(t)) - H_m(t)\gamma_m(t) \quad (64)$
  Subject to: $\quad 0 \leq \gamma_m(t) \leq A_m^{max}$

  This is a maximization of a concave single variable function over an interval, the same as in the internet algorithm of Section IV.

- (Flow Control) For each slot t, each session m observes $A_m(t)$ and the queue values $H_m(t)$, $Q_{n_m}^{(c_m)}(t)$ (where n_m denotes the source node of session m, and c_m represents its destination). Note that these queues are all local to the source node of the session, and hence can be observed easily. It then chooses $x_m(t)$ to solve:

  Maximize: $\quad H_m(t)x_m(t) - Q_{n_m}^{(c_m)}(t)x_m(t) \quad (65)$
  Subject to: $\quad 0 \leq x_m(t) \leq A_m(t)$

  This reduces to the "bang-bang" flow control decision of choosing $x_m(t) = A_m(t)$ if $Q_{n_m}^{(c_m)}(t) \leq H_m(t)$, and $x_m(t) = 0$ otherwise.

- (Resource Allocation and Transmission) For each slot t, the network controller observes queue backlogs $\{Q_n^{(c)}(t)\}$ and the topology state $S(t)$ and chooses $I(t) \in \mathcal{I}_{S(t)}$ and $\{\mu_{ij}^{(c)}(t)\}$ to solve:

  Max: $\sum_{n,c} Q_n^{(c)}(t)[\sum_{j=1}^{N} \mu_{nj}^{(c)}(t) - \sum_{i=1}^{N} \mu_{in}^{(c)}(t)]$ (66)
  S.t.: $\quad I(t) \in \mathcal{I}_{S(t)}$ and (57)-(59)

- (Queue Updates) Update the virtual queues $H_m(t)$ according to (63) and the acutal queues $Q_n^{(c)}(t)$ according to (56).

The exact decisions required to implement the resource allocation and transmission component are described in Subsection V-D below. Before covering this, we state the performance of the algorithm under a general C-approximate implementation of the above algorithm. For simplicity, assume all queues are initially zero. By Theorem 1c we have for any integers $R > 0$, $T > 0$:

$$
\begin{aligned}
\phi(\overline{x}) \geq \ & \frac{1}{R}\sum_{r=0}^{R-1}\Phi_r^* - \frac{B+C}{V} - \frac{D(T-1)}{V} \\
& - \sum_{m=1}^{M}\frac{\nu_m|H_m(RT)|}{RT}
\end{aligned}
\tag{67}
$$

While Theorem 1 also provides a bound on the final term, and bounds on all queue sizes, we can again provide tighter constant queue bounds by taking advantage of the flow control structure of the problem. Indeed, by the same argument that proves (53) in Section IV, we have that $H_m(t)$ cannot decrease if it is already negative, and cannot increase if it is beyond $V\nu_m$, so that for all $m \in \{1, \ldots, M\}$ we have:

$$
-A_m^{max} \leq H_m(t) \leq V\nu_m + A_m^{max} \quad \forall t
\tag{68}
$$

provided that these bounds are true for $H_m(0)$ (which is indeed the case if $H_m(0) = 0$). Therefore, $|H_m(t)| \leq V\nu_m + A_m^{max}$, and the utility bound (67) becomes:

$$
\begin{aligned}
\phi(\overline{x}) \geq \ & \frac{1}{R}\sum_{r=0}^{R-1}\Phi_r^* - \frac{B+C}{V} - \frac{D(T-1)}{V} \\
& - \sum_{m=1}^{M}\frac{\nu_m(V\nu_m + A_m^{max})}{RT}
\end{aligned}
\tag{69}
$$

The values of B and D in (69) for this context are given in Appendix E. Note that this yields a "utility fudge factor" of the form as indicated in the introduction of this paper:

$$
\text{utility fudge factor} = \frac{B_1 T}{V} + \frac{B_2 V}{RT}
$$

where:

$$
B_1 \triangleq (B + C - D)/T + D \ , \ B_2 \triangleq \sum_{m=1}^{M}\nu_m(\nu_m + A_m^{max}/V)
$$

The value of C used in the above bound is equal to 0 if we use a 0-approximation, being an exact implementation of the above algorithm. In the next subsections we purposefully engineer a C-approximation for a nonzero but constant C so that we can additionally provide deterministic bounds on all *actual* queues $Q_n^{(c)}(t)$.

D. Resource Allocation and Transmission

By switching the sums in (66), it is easy to show that the resource allocation and transmission maximization reduces to the following generalized "max-weight" algorithm (see [1]): Every slot t, choose $I(t) \in \mathcal{I}_{S(t)}$ to maximize:

$$
\sum_{i=1}^{N}\sum_{j=1}^{N}C_{ij}(I(t), S(t))W_{ij}(t)
$$

where $W_{ij}(t)$ are weights defined by:

$$
W_{ij}(t) \triangleq \max_{c\in\{1,\ldots,N\}}\max[W_{ij}^{(c)}(t), 0]
$$

where $W_{ij}^{(c)}(t)$ are differential backlogs:

$$
W_{ij}^{(c)}(t) \triangleq Q_i^{(c)}(t) - Q_j^{(c)}(t)
$$

The transmission decision variables are then given by:

$$
\mu_{ij}^{(c)}(t) = \begin{cases} C_{ij}(I(t), S(t)) & \text{if } c = c_{ij}^*(t) \text{ and } W_{ij}^{(c)}(t) \geq 0 \\ 0 & \text{otherwise} \end{cases}
$$

where $c_{ij}^*(t)$ is defined as the commodity $c \in \{1, \ldots, N\}$ that maximizes the differential backlog $W_{ij}^{(c)}(t)$ (breaking ties arbitrarily).

E. A C-Approximate Transmission Algorithm

Rather than implement the exact transmission algorithm in the above subsection, we present here a useful C-approximation that yields bounded queues (see also [9] [32]). Define $\hat{W}_{ij}^{(c)}(t)$ as follows:

$$
\hat{W}_{ij}^{(c)}(t) \triangleq
\begin{cases} W_{ij}^{(c)}(t) + \theta_i^{(c)} - \theta_j^{(c)} & \text{if } Q_j^{(c)}(t) \leq Q^{max} - \beta_j \\ -1 & \text{otherwise} \end{cases}
\tag{70}
$$

where for each $n \in \{1, \ldots, N\}$, β_n is defined as the largest amount of any commodity that can enter node n, considering both exogenous and endogenous arrivals (this is finite by the boundedness assumptions on transmission rates and new arrivals), and where Q^{max} is defined:

$$
Q^{max} \triangleq V\nu^{max} + A^{max} + \beta^{max}
\tag{71}
$$

where ν^{max}, A^{max}, and β^{max} are given by:

$$
\begin{aligned}
\nu^{max} &\triangleq \max_{m\in\{1,\ldots,M\}}\nu_m \\
A^{max} &\triangleq \max_{m\in\{1,\ldots,M\}}A_m^{max} \\
\beta^{max} &\triangleq \max_{n\in\{1,\ldots,N\}}\beta_n
\end{aligned}
$$

Finally, the values $\theta_i^{(c)}$ are any non-negative weights that represent some type of estimate of the distance from node i to destination c (possibly being zero if there is no such estimate available). Such weights are known to experimentally improve delay by biasing routing decisions to move in directions closer to the destination (see [6][1][9]). Then define $\hat{W}_{ij}(t)$ as:

$$
\hat{W}_{ij}(t) \triangleq \max_{c\in\{1,\ldots,N\}}\max[\hat{W}_{ij}^{(c)}(t), 0]
$$

and choose $I(t) \in \mathcal{I}_{\omega(t)}$ to maximize:

$$
\sum_{i=1}^{N}\sum_{j=1}^{N}C_{ij}(I(t), S(t))\hat{W}_{ij}(t)
\tag{72}
$$

and choose transmission variables:

$$
\mu_{ij}^{(c)}(t) = \begin{cases} C_{ij}(I(t), S(t)) & \text{if } c = \hat{c}_{ij}^*(t) \text{ and } \hat{W}_{ij}^{(c)}(t) \geq 0 \\ 0 & \text{otherwise} \end{cases}
\tag{73}
$$

where $\hat{c}^*(t)$ is the commodity $c \in \{1, \ldots, N\}$ that maximizes $\hat{W}_{ij}^{(c)}(t)$.

F. Bounded Queues

Lemma 3: (Bounded $Q_n^{(c)}(t)$) Suppose auxiliary variables and flow control decisions are made according to (64) and (65), with update equations (63) and (56). Suppose that $I(t) \in \mathcal{I}_{S(t)}$ is chosen in some arbitrary manner every slot t (not necessarily according to (72)), but that transmission decisions are made according to (73) with respect to the particular $I(t)$ chosen. Then for all t we have:

$$Q_n^{(c)}(t) \le Q^{max} = V\nu^{max} + A^{max} + \beta^{max} \qquad (74)$$

provided that this inequality holds at $t = 0$.

Proof: Suppose that $Q_n^{(c)}(t) \le Q^{max}$ for all n, c for a particular slot t (this is true by assumption on slot $t = 0$). We prove it also holds for slot $t + 1$. First suppose that $Q_n^{(c)}(t) \le Q^{max} - \beta_n$. Then, because β_n is the largest amount of new arrivals to queue $Q_n^{(c)}(t)$ over one slot (considering both endogenous and exogenous arrivals), it must be that $Q_n^{(c)}(t+1) \le Q^{max}$.

Consider now the opposite case when $Q^{max} - \beta_n < Q_n^{(c)}(t) \le Q^{max}$. Then from (70) we see that $\hat{W}_{in}^{(c)} = -1$ for all links (i, n) over which new commodity c data could be transmitted to node n from other nodes. Thus, by (73) we see that no commodity c data will be transmitted to node n from any other node. Further, We have:

$$Q_n^{(c)}(t) > Q^{max} - \beta_n = V\nu^{max} + A^{max} \ge H_m(t)$$

for all $m \in \{1, \ldots, M\}$, where the first equality follows by definition of Q^{max} in (71) and the final inequality follows by (68). It follows by the flow control decision (65) that $x_m(t) = 0$ for all sessions m that might deliver new data to queue $Q_n^{(c)}(t)$. Thus, no new commodity c data (exogenous or endogenous) arrives to node n on slot t, and $Q_n^{(c)}(t+1) \le Q_n^{(c)}(t) \le Q^{max}$. $\square$

The queue bound (74) in the above lemma provides the strong deterministic guarantee that all queues are bounded by a constant that grows linearly with the V parameter. Thus, while increasing V can improve the terms $(B + C)/V$ and $D(T - 1)/V$ in the utility guarantee (69), a tradeoff is in the linear growth with V in queue congestion (74), as well as the increase in the number of frames R required for the final term in the utility bound (69) to decay to near-zero.

While it is intuitive that the above algorithm produces a C-approximation for some constant value C, we complete the analysis below by formally showing this. Additionally, we note that a θ-multiplicative-approximate solution to (72) leads to utility guarantees where Φ_r^* is re-defined as a T-slot lookahead utility on a network where the $\boldsymbol{C}(I(t), S(t))$ function is replaced by $\theta\boldsymbol{C}(I(t), S(t))$, which holds for the same reason as described in Section IV-B.

G. Computing the C value

Lemma 4: Using the modified weights $\hat{W}_{ij}^{(c)}(t)$ in (70) results in a C-approximation of the max-weight resource allocation and transmission scheduling problem (66), with:

$$C \triangleq 2C_{sum}[\beta^{max} + \theta_{diff}] \qquad (75)$$

where C_{sum} is the largest possible sum of transmission rates $\sum_{ij} C_{ij}(I(t), S(t))$, summed over all links and considering all possible $S(t)$ states and $I(t)$ decisions (being finite by the boundedness assumptions on all links), and θ_{diff} is the maximum difference in $\theta_i^{(c)}$ and $\theta_j^{(c)}$, maximized over all node pairs (i, j) and all commodities c.

Proof: Because all queues $Q_n^{(c)}(t)$ are upper bounded by Q^{max}, if $Q_j^{(c)}(t) > Q^{max} - \beta_j$, then $\max[W_{ij}^{(c)}(t), 0] = \max[Q_i^{(c)}(t) - Q_j^{(c)}(t), 0] \le \beta_j$. It follows that:

$$|\max[W_{ij}^{(c)}(t), 0] - \max[\hat{W}_{ij}^{(c)}(t), 0]| \le \beta^{max} + \theta_{diff}$$

It follows that:

$$|W_{ij}(t) - \hat{W}_{ij}(t)| \le \beta^{max} + \theta_{diff}$$

Therefore:

$$|\sum_{i=1}^{N}\sum_{j=1}^{N} C_{ij}(I(t), S(t))[W_{ij}(t) - \hat{W}_{ij}(t)]|$$
$$\le \sum_{i=1}^{N}\sum_{j=1}^{N} C_{ij}(I(t), S(t))[\beta^{max} + \theta_{diff}]$$
$$\le C_{sum}[\beta^{max} + \theta_{diff}] = C/2$$

where C_{sum} is the maximum sum rate over all links on any slot. Now let $I^*(t)$ be the maximum of $\sum_{ij} C_{ij}(I(t), S(t))W_{ij}(t)$ over $\mathcal{I}_{S(t)}$, and let $\hat{I}(t)$ be the maximum of $\sum_{ij} C_{ij}(I(t), S(t))\hat{W}_{ij}(t)$. Then:

$$\sum_{ij} C_{ij}(\hat{I}(t), S(t))W_{ij}(t)$$
$$\ge \sum_{ij} C_{ij}(\hat{I}(t), S(t))\hat{W}_{ij}(t) - C/2$$
$$\ge \sum_{ij} C_{ij}(I^*(t), S(t))\hat{W}_{ij}(t) - C/2$$
$$\ge \sum_{ij} C_{ij}(I^*(t), S(t))W_{ij}(t) - C$$

It follows that the resource allocation $\hat{I}(t)$ (and the corresponding transmission decisions given by (73)) yields a C-approximation. $\square$

VI. Approximate Scheduling and Slater Conditions

Here we replace Assumption A2 with a stronger assumption that states the constraints can be satisfied with δ slackness. This is related to a *Slater condition* in classical static optimization problems [33]. It allows all queues to be deterministically bounded. It also allows performance analysis for implementations when the error in the attempted minimization of the right hand side of (38) is off by more than just a constant C, such as an amount that may be proportional to the queue backlog (similar to the θ-multiplicative-approximations discussed in Section IV-B).

Assumption A3: There exists a value $\delta > 0$ such that for all $\omega \in \{\omega(0), \ldots, \omega(t_{end} - 1)\}$, there is at least one control

action $\alpha'_\omega \in \mathcal{A}_\omega$ that satisfies:

$$\hat{y}_l(\alpha'_\omega, \omega) + g_l(\hat{\boldsymbol{x}}(\alpha'_\omega, \omega)) \leq -\delta \ \forall l \in \{1, \ldots, L\}$$
$$\hat{a}_k(\alpha'_\omega, \omega) \leq \hat{b}_k(\alpha'_\omega, \omega) - \delta \ \forall k \in \{1, \ldots, K\}$$
$$\hat{\boldsymbol{x}}(\alpha'_\omega, \omega) + \boldsymbol{\epsilon} \in \mathcal{X}$$

for all vectors $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_M)$ with entries ϵ_m that satisfy $|\epsilon_m| \leq \delta$ for all $m \in \{1, \ldots, M\}$. Further, assume that:

$$x_m^{min} + \delta \leq \hat{x}_m(\alpha'_\omega, \omega) \leq x_m^{max} - \delta \ \forall m \in \{1, \ldots, M\}$$

This final assumption is mild and can easily be engineered to be true by convexly extending the range of the convex functions $f(\cdot)$, $g_l(\cdot)$ by δ in all directions (so that x_m^{min} is decreased by δ and x_m^{max} is increased by δ), as in [34].

Define a $C(t)$-*approximation* as an algorithm that, every slot t, observes the current queue states and choses a control action that comes within $C(t)$ of minimizing the right hand side of (38), where $C(t)$ is a value that can depend on t. Suppose that we implement the universal scheduling algorithm of Section III-B using a $C(t)$-approximation with $C(t)$ that satisfies the following for all t:

$$C(t) \leq C + V\epsilon_V + \sum_{l=1}^{L} Z_l(t)\epsilon_Z$$
$$+ \sum_{k=1}^{K} Q_k(T)\epsilon_Q + \sum_{m=1}^{M} |H_m(t)|\epsilon_H \quad (76)$$

where $C, \epsilon_V, \epsilon_Z, \epsilon_Q, \epsilon_H$ are non-negative constants. Note that this is a C-approximation if $\epsilon_V = \epsilon_Z = \epsilon_Q = \epsilon_H = 0$, and is the exact minimization of (38) if we additionally have $C = 0$.

Theorem 2: Suppose Assumptions A1 and A3 hold for some $\delta > 0$. Consider any $C(t)$-approximate algorithm that satisfies (76) every slot t, and assume that:

$$\epsilon_Q < \delta \ , \ \epsilon_H < \delta$$
$$\epsilon_Z + \epsilon_H \sum_{m=1}^{M} \beta_{l,m} < \delta \ \forall l \in \{1, \ldots, L\}$$

Let the random event sequence $\{\omega(0), \omega(1), \ldots\}$ be arbitrary. Suppose all initial queue backlogs are zero. Then:

(a) All queue backlogs are bounded, so that for any slot $t \geq 0$ we have:

$$Q_k(t), Z_l(t), |H_m(t)| \leq \frac{VC_3}{\theta}$$

where

$$C_3 \triangleq \sqrt{D_1 + D_2 + D_3}$$

where D_1, D_2, D_3 are constants defined as:

$$D_1 \triangleq \left[\frac{B+C}{V} + (y_0^{max} - y_0^{min}) + (f^{max} - f^{min}) + \epsilon_V\right]^2$$
$$D_2 \triangleq 2D\theta^2/V^2$$
$$D_3 \triangleq \frac{2z_{max}\theta}{V}\sqrt{D_1}$$

where D is defined in (41), z_{max} is the maximum over all z_l^{diff}, q_k^{diff}, and h_m^{diff} constants, and where θ is defined:

$$\theta \triangleq \min\left[\delta - \epsilon_Q, \delta - \epsilon_H, \frac{\delta - \epsilon_Z - \epsilon_H\beta_{sum}}{1 + \beta_{sum}}\right] \quad (77)$$

where β_{sum} is defined:

$$\beta_{sum} \triangleq \max_{l \in \{1, \ldots, L\}} \sum_{m=1}^{M} \beta_{l,m}$$

In the special case when $\epsilon_Z = \epsilon_Q = \epsilon_H = 0$, we have $\theta = \delta/(1 + \beta_{sum})$.

(b) For any designated time $t_{end} > 0$ we have:

$$\overline{y}_l + g_l(\overline{\boldsymbol{x}}) \leq \frac{VC_3}{\theta t_{end}}\left[1 + \sum_{m=1}^{M} \beta_{l,m}\right] \quad \forall l \in \{1, \ldots, L\}$$
$$\overline{a}_k \leq \overline{b}_k + \frac{VC_3}{\theta t_{end}} \quad \forall k \in \{1, \ldots, K\}$$
$$\overline{\boldsymbol{x}} + \overline{\boldsymbol{\epsilon}}(t) \in \mathcal{X}$$

where $\boldsymbol{\epsilon}(t) = (\epsilon_1(t), \ldots, \epsilon_M(t))$ has entries that satisfy:

$$|\epsilon_m(t)| \leq \frac{VC_3}{\theta t_{end}}$$

c) Consider any positive integer frame size T, any positive integer R, and define $t_{end} = RT$. Then the value of the system cost metric over t_{end} slots satisfies:

$$\overline{y}_0 + f(\overline{\boldsymbol{x}}) \leq \frac{(1-p)}{R}\sum_{r=0}^{R-1} F_r^* + \epsilon_V + p(y_0^{max} + f^{max})$$
$$+ \frac{B + C + \tilde{D}(T-1)}{V} + \sum_{m=1}^{M} \frac{\nu_m VC_3}{\theta RT} \quad (78)$$

where p is defined:

$$p \triangleq \max\left[\frac{\epsilon_Z}{\delta - (\epsilon_H + \theta)\beta_{sum}}, \frac{\epsilon_Q}{\delta}, \frac{\epsilon_H}{\epsilon_H + \theta}\right]$$

and where $\tilde{D}$ is a constant defined in (89), B is a constant defined in (34), and C_3 is a constant defined in part (a).

Proof: See Appendix D. $\square$

The cost bound (78) can be understood as follows: The last term on the right hand side goes to zero as R increases (and is equal to 0 for all R if $f(\cdot) = 0$ so that $\nu_m = 0$ for all m). The second to last term can be made arbitrarily small with a suitably large V. Finally, if $\epsilon_Z, \epsilon_Q, \epsilon_H, \epsilon_V$ are small, then p is small and the remaining terms on the right hand side are close to the cost $\frac{1}{R}\sum_{r=0}^{R-1} F_r^*$ which is associated with implementing the T-slot lookahead policy over R frames.

VII. CONCLUSIONS

We have developed a framework for universal constrained optimization of time averages in time varying systems. Our results hold for any event sample paths and do not require a probability model. It was shown that performance can closely track the performance of an ideal policy with knowledge of the future up to T slots, provided that we allow the number of T-slot frames, denoted by R, to be large enough to ensure that the error terms decay to a negligible value. This framework was applied to an internet model and to a more extensive queueing network model to provide utility guarantees with deterministic queue bounds for arbitrary traffic, channels, and mobility.

APPENDIX A — PROOF OF THEOREM 1

We first prove parts (a) and (b) of Theorem 1.

Proof: (Theorem 1 part (a)) Let $\boldsymbol{\gamma}(t)$ and $\alpha(t)$ represent the decisions made by the C-approximate policy on slot t, which necessarily satisfy constraints (21)-(23). Because these decisions come within C of minimizing the right hand side of (38) over all other possible decisions, we have from (38):

$$\Delta_1(t) + V\hat{y}_0(\alpha(t), \omega(t)) + Vf(\boldsymbol{\gamma}(t)) \leq$$
$$B + C + V\hat{y}_0(\alpha^*(t), \omega(t)) + Vf(\boldsymbol{\gamma}^*(t))$$
$$+ \sum_{l=1}^{L} Z_l(t)[\hat{y}_l(\alpha^*(t), \omega(t)) + g_l(\boldsymbol{\gamma}^*(t))]$$
$$+ \sum_{k=1}^{K} Q_k(t)[\hat{a}_k(\alpha^*(t), \omega(t)) - \hat{b}_k(\alpha^*(t), \omega(t))]$$
$$+ \sum_{m=1}^{M} H_m(t)[\gamma_m^*(t) - \hat{x}_m(\alpha^*(t), \omega(t))] \quad (79)$$

where $\alpha^*(t)$ and $\boldsymbol{\gamma}^*(t)$ represent any alternative decisions that could be made on slot t that satisfy (21)-(23).

Now choose $\alpha^*(t) = \alpha'_{\omega(t)}$, where $\alpha'_{\omega(t)} \in \mathcal{A}_{\omega(t)}$ is the decision known to exist by Assumption A2 that satisfies:

$$\hat{y}_l(\alpha^*(t), \omega(t)) + g_l(\hat{\boldsymbol{x}}(\alpha^*(t), \omega(t))) \leq 0 \quad \forall l \in \{1, \ldots, L\}$$
$$\hat{a}_k(\alpha^*(t), \omega(t)) \leq \hat{b}_k(\alpha^*(t), \omega(t)) \quad \forall k \in \{1, \ldots, K\}$$
$$\hat{\boldsymbol{x}}(\alpha^*(t), \omega(t)) \in \mathcal{X}$$

Further, choose $\boldsymbol{\gamma}^*(t) = \hat{\boldsymbol{x}}(\alpha^*(t), \omega(t))$. These decisions satisfy (21)-(23), and plugging these decisions directly into the right hand side of (79) yields:

$$\Delta_1(t) + V\hat{y}_0(\alpha(t), \omega(t)) + Vf(\boldsymbol{\gamma}(t)) \leq$$
$$B + C + V\hat{y}_0(\alpha^*(t), \omega(t)) + Vf(\boldsymbol{\gamma}^*(t))$$

Rearranging terms and using the bounds y_0^{min}, y_0^{max} and f^{min}, f^{max} yields:

$$\Delta_1(t) \leq B + C + V(y_0^{max} - y_0^{min}) + V(f^{max} - f^{min})$$

Let the right hand side of the above inequality be denoted by P. Using the definition of $\Delta_1(t)$ thus gives:

$$L(\boldsymbol{\Theta}(t+1)) - L(\boldsymbol{\Theta}(t)) \leq P$$

The above holds for all $t \geq 0$. Summing over $\tau \in \{0, \ldots, t-1\}$ (for some time $t > 0$) and dividing by t yields:

$$\frac{1}{t}[L(\boldsymbol{\Theta}(t)) - L(\boldsymbol{\Theta}(0))] \leq P$$

Using the definition of $L(\boldsymbol{\Theta}(t))$ in (32) proves part (a). □

Proof: (Theorem 1 part (b)) From part (a), if all queues are initially empty (so that $L(\boldsymbol{\Theta}(0)) = 0$), we have for all slots $t > 0$:

$$Z_l(t), Q_k(t), |H_m(t)| \leq C_0\sqrt{tV} \quad (80)$$

Plugging (80) into (26), (27), (28) of Lemma 1 proves part (b). □

To prove part (c) of Theorem 1, we need the following preliminary lemma.

Lemma 5: For any initial time t_0, any queue values $\boldsymbol{\Theta}(t_0)$, any integer $T > 1$, and any collection of C-approximate decisions that are implemented over the T-slot interval $\tau \in \{t_0, \ldots, t_0 + T - 1\}$, we have:

$$\Delta_T(t_0) + \sum_{\tau=t_0}^{t_0+T-1} [V\hat{y}_0(\alpha(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}(\tau))] \leq$$
$$BT + CT + DT(T-1)$$
$$+ \sum_{\tau=t_0}^{t_0+T-1} [V\hat{y}_0(\alpha^*(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}^*(\tau))]$$
$$+ \sum_{l=1}^{L} Z_l(t_0) \sum_{\tau=t_0}^{t_0+T-1} [\hat{y}_l(\alpha^*(\tau), \omega(\tau)) + g_l(\boldsymbol{\gamma}^*(\tau))]$$
$$+ \sum_{k=1}^{K} Q_k(t_0) \sum_{\tau=t_0}^{t_0+T-1} [\hat{a}_k(\alpha^*(\tau), \omega(\tau)) - \hat{b}_k(\alpha^*(\tau), \omega(\tau))]$$
$$+ \sum_{m=1}^{M} H_m(t_0) \sum_{\tau=t_0}^{t_0+T-1} [\gamma_m^*(\tau) - \hat{x}_m(\alpha^*(\tau), \omega(\tau))]$$

for any alternative decisions $\alpha^*(\tau)$, $\boldsymbol{\gamma}^*(\tau)$ over $\tau \in \{t_0, \ldots, t_0 + T - 1\}$ that satisfy (21)-(23). The constant B is defined according to (34) and D is defined by (41).

Proof: (Lemma 5) Because our policy is C-approximate, for all slots t it comes within C of minimizing the right hand side of (38). Hence, for all $\tau \in \{t_0, \ldots, t_0 + T - 1\}$ we have:

$$\Delta_1(\tau) + V\hat{y}_0(\alpha(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}(\tau)) \leq$$
$$B + C + V\hat{y}_0(\alpha^*(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}^*(\tau))$$
$$+ \sum_{l=1}^{L} Z_l(\tau)[\hat{y}_l(\alpha^*(\tau), \omega(\tau)) + g_l(\boldsymbol{\gamma}^*(\tau))]$$
$$+ \sum_{k=1}^{K} Q_k(\tau)[\hat{a}_k(\alpha^*(\tau), \omega(\tau)) - \hat{b}_k(\alpha^*(\tau), \omega(\tau))]$$
$$+ \sum_{m=1}^{M} H_m(\tau)[\gamma_m^*(\tau) - \hat{x}_m(\alpha^*(\tau), \omega(\tau))] \quad (81)$$

However, by definition of z_l^{diff}, q_k^{diff}, h_m^{diff}, the queues $Z_l(t)$, $Q_k(t)$, $H_m(t)$ can change by at most these values on each slot, and hence for $\tau \in \{t_0, \ldots, t_0 + T - 1\}$ we have:

$$|Z_l(\tau) - Z_l(t_0)| \leq z_l^{diff} \cdot (\tau - t_0)$$
$$|Q_k(\tau) - Q_k(t_0)| \leq q_k^{diff} \cdot (\tau - t_0)$$
$$|H_m(\tau) - H_m(t_0)| \leq h_m^{diff} \cdot (\tau - t_0)$$

Using these in (81) gives:

$$\Delta_1(\tau) + V\hat{y}_0(\alpha(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}(\tau)) \leq$$
$$B + C + 2D \cdot (\tau - t_0) + V\hat{y}_0(\alpha^*(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}^*(\tau))$$
$$+ \sum_{l=1}^{L} Z_l(t_0)[\hat{y}_l(\alpha^*(\tau), \omega(\tau)) + g_l(\boldsymbol{\gamma}^*(\tau))]$$
$$+ \sum_{k=1}^{K} Q_k(t_0)[\hat{a}_k(\alpha^*(\tau), \omega(\tau)) - \hat{b}_k(\alpha^*(\tau), \omega(\tau))]$$
$$+ \sum_{m=1}^{M} H_m(t_0)[\gamma_m^*(\tau) - \hat{x}_m(\alpha^*(\tau), \omega(\tau))]$$

where D is defined in (41). Summing the above inequality over $\tau \in \{t_0, \ldots, t_0 + T - 1\}$ yields the result, where we use the fact that:

$$\sum_{\tau=t_0}^{t_0+T-1} (\tau - t_0) = T(T-1)/2$$

$\square$

We can now prove Theorem 1 part (c).

Proof: (Theorem 1 part (c)) Fix integers $r \geq 0$ and $T > 0$. Fix $\epsilon > 0$, and let $\alpha^*(\tau)$ represent the decisions over the interval $\tau \in \{rT, \ldots, (r+1)T - 1\}$ that solve the problem (15) and yield cost that is no more than $F_r^* + \epsilon$. Let $\boldsymbol{\gamma}^*(\tau)$ be constant over $\tau \in \{rT, \ldots, (r+1)T - 1\}$, given by:

$$\boldsymbol{\gamma}^*(\tau) = \frac{1}{T} \sum_{t=rT}^{(r+1)T-1} \hat{\boldsymbol{x}}(\alpha^*(t), \omega(t))$$

Plugging these alternative decisions $\alpha^*(\tau)$ and $\boldsymbol{\gamma}^*(\tau)$ into the result of Lemma 5 for $t_0 = rT$ yields:

$$\Delta_T(rT) + \sum_{\tau=rT}^{(r+1)T-1} [V\hat{y}_0(\alpha(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}(\tau))] \leq$$
$$BT + CT + DT(T-1) + VT(F_r^* + \epsilon)$$

The above holds for all $\epsilon > 0$, and hence we can take a limit as $\epsilon \to 0$ to remove the ϵ in the final term. Define $t_{end} = RT$ for some positive integer R. Summing the above over $r \in \{0, \ldots, R-1\}$ and dividing by VRT yields:

$$\overline{y}_0 + f(\overline{\boldsymbol{\gamma}}) + \frac{L(\boldsymbol{\Theta}(RT)) - L(\boldsymbol{\Theta}(0))}{VRT} \leq$$
$$\frac{1}{R} \sum_{r=0}^{R-1} F_r^* + \frac{B+C}{V} + \frac{D(T-1)}{V} \qquad (82)$$

where $\overline{y}_0$ and $\overline{\boldsymbol{\gamma}}$ represent time averages over the first t_{end} slots, and where we have used Jensen's inequality in the concave function $f(\boldsymbol{\gamma})$.

However, we have by (12):

$$f(\overline{\boldsymbol{x}}) \leq f(\overline{\boldsymbol{\gamma}}) + \sum_{m=1}^{M} \nu_m |\overline{\gamma}_m - \overline{x}_m|$$
$$= f(\overline{\boldsymbol{\gamma}}) + \sum_{m=1}^{M} \frac{\nu_m |H_m(RT) - H_m(0)|}{RT}$$

where the final equality holds by (29). Using this in (82) together with the fact that $L(\cdot) \geq 0$ yields the cost bound (40) of part (c). Finally, if initial queue backlogs are 0, by (80) applied to time $t = RT$ we have:

$$\frac{|H_m(RT) - H_m(0)|}{RT} = \frac{|H_m(RT)|}{RT} \leq \frac{C_0 \sqrt{V}}{\sqrt{RT}}$$

$\square$

APPENDIX B — CONDITIONS FOR ACHIEVABILITY OF F^*

Consider the following additional assumption.

Assumption A4: We have either one of the following two conditions:

1) For all $\omega \in \{\omega(0), \ldots, \omega(t_{end} - 1)\}$, the control action space $\mathcal{A}_\omega$ contains a finite number of actions.
2) For all $\omega \in \{\omega(0), \ldots, \omega(t_{end} - 1)\}$, the set $\mathcal{A}_\omega$ is a compact subset of $\mathbb{R}^c$ for some dimension c, the functions $\hat{y}_l(\alpha, \omega), \hat{a}_k(\alpha, \omega)$ are lower semi-continuous over $\alpha \in \mathcal{A}_\omega$, the functions $\hat{b}_k(\alpha, \omega)$ are upper semi-continuous over $\alpha \in \mathcal{A}_\omega$, and the functions $\hat{x}_m(\alpha, \omega)$ are continuous over $\alpha \in \mathcal{A}_\omega$. Note that all continuous functions are both upper and lower semi-continuous.[8]

Lemma 6: Suppose Assumptions A1, A2, A4 hold for given values $\{\omega(0), \ldots, \omega(t_{end} - 1)\}$. Then the infimum value F^* for the problem (3)-(7) can be achieved by a particular (possibly non-unique) sequence of control actions $\{\alpha^*(0), \ldots, \alpha^*(t_{end} - 1)\}$. That is, these actions satisfy the feasibility constraints (4)-(7), and yield:

$$\overline{y}_0 + f(\overline{x}_1, \ldots, \overline{x}_M) = F^*$$

where:

$$\overline{x}_m = \frac{1}{t_{end}} \sum_{\tau=0}^{t_{end}-1} \hat{x}_m(\alpha^*(\tau), \omega(\tau)) \quad \forall m \in \{1, \ldots, M\}$$

and where $\overline{y}_0$ is similarly defined as a time average over $\tau \in \{0, \ldots, t_{end} - 1\}$.

Proof: We already know that Assumption A2 implies the existence of a feasible sequence of control actions. Thus, the infimum value F^* of the cost metric over all feasible policies is well defined, and by Assumption A1 it must satisfy:

$$y_0^{min} + f^{min} \leq F^* \leq y_0^{max} + f^{max}$$

Consider now the case when the first condition of Assumption A4 holds. Then there are only a finite number of possible control sequences over the horizon $\{0, 1, \ldots, t_{end} - 1\}$, and so there is one that achieves the minimum cost value F^*. The case when the second condition of Assumption A4 holds can be proven using the Bolzano-Wierstrass Theorem together with a simple limiting argument, and is omitted for brevity. $\square$

APPENDIX C – ERGODICITY

Consider the infinite horizon problem discussed in Section III-E. Suppose that the random events $\{\omega(0), \omega(1), \omega(2), \ldots\}$ evolve according to a general ergodic process with a well defined time average probability distribution. Specifically, let Ω represent a finite (but arbitrarily large) outcome space for $\omega(t)$, and for each $\omega \in \Omega$ assume that there is a steady state value $\pi(\omega)$, such that:

$$\lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} 1_\omega(\tau) = \pi(\omega) \quad \text{with probability 1}$$

where $1_\omega(\tau)$ is an indicator function that is 1 if $\omega(\tau) = \omega$, and zero else. Further, assume the limiting probability converges

---

[8] A function $b(\boldsymbol{\alpha})$ is *upper semi-continuous* over $\boldsymbol{\alpha} \in \mathcal{A}$ if for any $\boldsymbol{\alpha} \in \mathcal{A}$, we have $b(\boldsymbol{\alpha}) \geq \lim_{n \to \infty} b(\boldsymbol{\beta}_n)$ for all sequences $\boldsymbol{\beta}_n \in \mathcal{A}$ such that $\lim_{n \to \infty} \boldsymbol{\beta}_n = \boldsymbol{\alpha}$. A function is *lower semi-continuous* if the inequality is reversed. All bounded functions that are discontinuous only on a set of measure zero can be easily modified to have the desired semi-continuous property by appropriately re-defining the function value at points of discontinuity. Most systems of practical interest have the desired semi-continuity properties.

uniformly to the steady state value, regardless of past history, so that:

$$|Pr[\omega(t+t_0) = \omega|History(t_0)] - \pi(\omega)| \le error(t)$$

where $History(t_0)$ represents the past history of the process up to slot t_0, and where $error(t)$ is a function that decays to 0 as $t \to \infty$, regardless of the past history. This is related to the *decaying memory property* in [15] and the admissibility assumptions in [4][1].

In this case, it can be shown that the optimal infinite horizon cost, denoted f^*, can be achieved over the class of stationary and randomized algorithms that make (possibly probabilistic) decisions for control actions on each slot t based only on the current state $\omega(t)$ (see [4][6][32] for related proofs).[9] Further, under mild conditions (such as the existence of a value $\delta > 0$ for which the Slater condition of Assumption A3 in Section VI is satisfied), the value of F_r^*, being the optimal time average cost under the T-slot lookahead policy over the T-slot interval starting at time rT, satisfies for any integer $r \ge 0$:

$$\lim_{T \to \infty} F_r^* = f^* \quad \text{with probability 1}$$

That is, regardless of the past history before time rT, the T-slot lookahead policy over a very large T approaches the optimal f^*. The reason the "mild" additional conditions, such as the Slater condition, is needed, is that F_r^* requires all constraints to be exactly satisfied by the end of the T slots, whereas the infinite horizon problem does not require this.

Because of the uniform error decay, we have:

$$|\mathbb{E}\{F_r^*\} - f^*| \le \delta(T)$$

where $\delta(T)$ is a function such that $\delta(T) \to 0$ as $T \to \infty$. Therefore we can write:

$$F_r^* = f^* + \delta_r$$

where δ_r is a random variable that satisfies $|\mathbb{E}\{\delta_r\}| \le \delta(T)$ for all r.

Using the definition of δ_r, it follows from (45) that time average cost satisfies for any integer $T > 0$:

$$\lim_{R \to \infty} [\overline{y}_0(RT) + f(\overline{x}(RT))] \le f^* + \frac{B+C}{V} + \frac{D(T-1)}{V}$$
$$+ \lim_{R \to \infty} \frac{1}{R} \sum_{r=0}^{R-1} \delta_r$$

Under mild conditions, such as when $\{\omega(t)\}$ evolves according to a finite state ergodic Markov chain, law of large number averaging principles imply that the last term, being a time average of the δ_r values, is bounded in absolute value with probability 1 by $\delta(T)$, a term that is negligibly small for large values of T. We can choose a large T provided that we also compensate with a large V to make the $D(T-1)/V$ term neglibible. This demonstrates that, with probability 1, the algorithm implemented over an infinite time horizon yields

[9]Similar results on optimality of stationary policies can typically be achieved when the cardinality of the set Ω is infinite, although steady state time averages and uniform convergence to a steady state are more awkward to deal with in this case. The easiest such arguments for (possibly uncountably) infinite sets Ω are for $\omega(t)$ processes that are i.i.d. over slots.

cost that can be pushed arbitrarily close to the optimal value f^* if V is suitably large.

APPENDIX D – PROOF OF THEOREM 2

Proof: (Theorem 2a) Define θ as the positive real number that solves the following problem (it can be shown that the solution is given by (77)):

$$\text{Maximize:} \qquad \theta \qquad (83)$$
$$\text{Subject to:} \qquad \theta \le \delta - \epsilon_Q$$
$$\theta \le \delta - \epsilon_Z - (\epsilon_H + \theta) \sum_{m=1}^{M} \beta_{l,m} \quad \forall l$$
$$\theta \le \delta - \epsilon_H$$

Following the proof of Theorem 1a and replacing C with $C(t)$ we have (compare with (79)):

$$\Delta_1(t) + V\hat{y}_0(\alpha(t), \omega(t)) + Vf(\gamma(t)) \le$$
$$B + C + V\hat{y}_0(\alpha^*(t), \omega(t)) + Vf(\gamma^*(t)) + V\epsilon_V$$
$$+ \sum_{l=1}^{L} Z_l(t)[\epsilon_Z + \hat{y}_l(\alpha^*(t), \omega(t)) + g_l(\gamma^*(t))]$$
$$+ \sum_{k=1}^{K} Q_k(t)[\epsilon_Q + \hat{b}_k(\alpha^*(t), \omega(t)) - \hat{a}_k(\alpha^*(t), \omega(t))]$$
$$+ \sum_{m=1}^{M} |H_m(t)|\epsilon_H$$
$$+ \sum_{m=1}^{M} H_m(t)[\gamma_m^*(t) - \hat{x}_m(\alpha^*(t), \omega(t))] \quad (84)$$

where $\alpha^*(t)$ and $\gamma^*(t)$ represent any alternative decisions that could be made on slot t that satisfy (21)-(23).

Now choose $\alpha^*(t) = \alpha'_{\omega(t)}$, where $\alpha'_{\omega(t)} \in \mathcal{A}_{\omega(t)}$ is the decision known to exist by Assumption A3. Further, choose $\gamma^*(t) = \gamma'(t)$, where $\gamma'(t) = (\gamma'_m(t))|_{m=1}^{M}$ is defined such that for all $m \in \{1, \dots, M\}$:

$$\gamma'_m(t) \triangleq \begin{cases} \hat{x}_m(\alpha'_{\omega(t)}, \omega(t)) - \epsilon_H - \theta & \text{if } H_m(t) \ge 0 \\ \hat{x}_m(\alpha'_{\omega(t)}, \omega(t)) + \epsilon_H + \theta & \text{if } H_m(t) < 0 \end{cases} \quad (85)$$

This is feasible because of the last inequality in Assumption A3 together with the fact that:

$$-\delta \le -\epsilon_H - \theta \le \epsilon_H + \theta \le \delta$$

With these choices, (84) becomes:

$$\Delta_1(t) \le$$
$$B + C + V(y_0^{max} - y_0^{min}) + V(f^{max} - f^{min}) + V\epsilon_V$$
$$- \sum_{l=1}^{L} Z_l(t)[\delta - \epsilon_Z - (\epsilon_H + \theta) \sum_{m=1}^{M} \beta_{l,m}]$$
$$- \sum_{k=1}^{K} Q_k(t)[\delta - \epsilon_Q]$$
$$- \sum_{m=1}^{M} |H_m(t)|\theta \quad (86)$$

where have used the fact that, from (13):

$$g_l(\boldsymbol{\gamma}^*(t)) \leq g_l(\hat{\boldsymbol{x}}(\alpha'_{\omega(t)}, \omega(t))) + (\epsilon_H + \theta) \sum_{m=1}^{M} \beta_{l,m}$$

Now define P as:

$$P \triangleq B + C + V(y_0^{max} - y_0^{min}) + V(f^{max} - f^{min}) + V\epsilon_V$$

Because the value θ is a bound on all the terms multiplying queue values in (86), we have:

$$\Delta_1(t) \leq P - \theta \sum_{l=1}^{L} Z_l(t) - \theta \sum_{k=1}^{K} Q_k(t) - \theta \sum_{m=1}^{M} |H_m(t)| \quad (87)$$

It follows that the drift is non-positive whenever the sum of the absolute value of queue size is greater than or equal to P/θ. It is not difficult to show that the largest possible value of $L(\boldsymbol{\Theta}(t))$ under the constraint that the sum of absolute queue values is less than or equal to P/θ is $(1/2)(P/\theta)^2$. Hence, if the Lyapunov function is larger than this value, it cannot increase on the next slot. However, if the absolute sum on slot t is less than or equal to P/θ we have:

$$\begin{aligned} L(\boldsymbol{\Theta}(t+1)) &\leq \frac{1}{2} \sum_{l=1}^{L} (Z_l(t) + z_l^{diff}(t))^2 \\ &\quad + \frac{1}{2} \sum_{k=1}^{K} (Q_k(t) + q_k^{diff}(t))^2 \\ &\quad + \frac{1}{2} \sum_{m=1}^{M} (|H_m(t)| + h_m^{diff}(t))^2 \\ &\leq L(\boldsymbol{\Theta}(t)) + D + z_{max} P/\theta \\ &\leq (1/2)(P/\theta)^2 + D + z_{max} P/\theta \quad (88) \end{aligned}$$

where $z_l^{diff}(t)$, $q_k^{diff}(t)$, $h_m^{diff}(t)$ represent the absolute value of the change in $Z_l(t)$, $Q_k(t)$, $H_m(t)$, respectively, over one slot, having maximum absolute value given by z_l^{diff}, q_k^{diff}, and h_m^{diff}, D is defined in (41), and z_{max} is the maximum over all z_l^{diff}, q_k^{diff}, and h_m^{diff} constants.

It follows that for all t we have:[10]

$$L(\boldsymbol{\Theta}(t)) \leq (1/2)(P/\theta)^2 + D + z_{max} P/\theta$$

Therefore, all queues are bounded by:

$$Q_k(t), Z_l(t), |H_m(t)| \leq \sqrt{(P/\theta)^2 + 2D + 2z_{max} P/\theta}$$

This bound is given by:

$$\frac{V\sqrt{P^2/V^2 + 2D\theta^2/V^2 + 2z_{max}\theta P/V^2}}{\theta}$$
$$= \frac{V\sqrt{D_1 + D_2 + D_3}}{\theta}$$

---

[10]More precisely, the bound on $L(\boldsymbol{\Theta}(t))$ is clearly true for $t = 0$. Supposing it is true for slot t, we show it is true for slot $t+1$: If the absolute sum is greater than or equal to P/θ on slot t, then the Lyapunov value cannot increase on the next slot and so the bound holds for slot $t+1$. Else, if the absolute sum is less than P/θ on slot t, then the bound again holds for slot $t+1$ by the calculation (88).

where

$$\begin{aligned} D_1 &\triangleq \left[\frac{B+C}{V} + (y_0^{max} - y_0^{min}) + (f^{max} - f^{min}) + \epsilon_V \right]^2 \\ D_2 &\triangleq 2D\theta^2/V^2 \\ D_3 &\triangleq \frac{2z_{max}\theta}{V}\sqrt{D_1} \end{aligned}$$

$\square$

Proof: (Theorem 2b) The proof follows immediately by applying the queue bounds of part (a) to the constraint bounds (26), (27), (28) of Lemma 1, using initial queue values of 0.

$\square$

Proof: (Theorem 2c) Fix integers $T > 0$ and $r \geq 0$. Similar to the proof of Lemma 5, we have by replacing C with $C(t)$ (compare with the bound in Lemma 5):

$$\begin{aligned} &\Delta_T(rT) + \sum_{\tau=rT}^{rT+T-1} [V\hat{y}_0(\alpha(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}(\tau))] \leq \\ &\qquad\qquad BT + CT + \tilde{D}T(T-1) \\ &\quad + \sum_{\tau=rT}^{rT+T-1} [V\epsilon_V + V\hat{y}_0(\alpha^*(\tau), \omega(\tau)) + Vf(\boldsymbol{\gamma}^*(\tau))] \\ &\quad + \sum_{l=1}^{L} Z_l(rT) \sum_{\tau=rT}^{rT+T-1} [\epsilon_Z + \hat{y}_l(\alpha^*(\tau), \omega(\tau)) + g_l(\boldsymbol{\gamma}^*(\tau))] \\ &\qquad\qquad + \sum_{k=1}^{K} Q_k(rT) \sum_{\tau=rT}^{rT+T-1} \epsilon_Q \\ &\quad + \sum_{k=1}^{K} Q_k(rT) \sum_{\tau=rT}^{rT+T-1} \left[\hat{a}_k(\alpha^*(\tau), \omega(\tau)) - \hat{b}_k(\alpha^*(\tau), \omega(\tau)) \right] \\ &\qquad\qquad + \sum_{m=1}^{M} |H_m(rT)|\epsilon_H \\ &\quad + \sum_{m=1}^{M} H_m(rT) \sum_{\tau=rT}^{rT+T-1} [\gamma_m^*(\tau) - \hat{x}_m(\alpha^*(\tau), \omega(\tau))] \end{aligned}$$

for any alternative decisions $\alpha^*(\tau)$, $\boldsymbol{\gamma}^*(\tau)$ over $\tau \in \{rT, \ldots, rT + T - 1\}$ that satisfy (21)-(23). The constant B is defined according to (34) and $\tilde{D}$ is defined by:

$$\tilde{D} \triangleq D + \frac{1}{2} \sum_{l=1}^{L} z_l^{diff} \epsilon_Z + \frac{1}{2} \sum_{k=1}^{K} q_k^{diff} \epsilon_Q + \frac{1}{2} \sum_{m=1}^{M} h_m^{diff} \epsilon_H$$

$$(89)$$

where D is defined in (41). Note that the above bound holds *deterministically* for all possible alternative (possibly randomized) policies. Hence, the bound also deterministically holds when the right hand side is replaced by the *expectation* over any particular randomized policy.[11]

Consider now the following randomized decisions for $\alpha^*(t)$ and $\boldsymbol{\gamma}^*(t)$: With probability p (to be defined later), for all slots $\tau \in \{rT, \ldots, rT + T - 1\}$, choose $\alpha^*(\tau) = \alpha'_{\omega(\tau)}$ and $\boldsymbol{\gamma}^*(\tau) = \boldsymbol{\gamma}'(\tau)$, where $\alpha'_{\omega(\tau)}$ and $\boldsymbol{\gamma}'(\tau)$ are the policies from the proof of part (a) associated with slot τ. Specifically, $\alpha'_{\omega(\tau)}$

---

[11]Formally, this uses the fact that if $b \leq \psi(\alpha_1, \ldots, \alpha_M)$ for all vectors $(\alpha_1, \ldots, \alpha_M) \in \mathcal{A}$ for some function $\psi(\cdot)$, and some constant b, then $b \leq \psi(A_1, \ldots, A_M)$ for any random vector $(A_1, \ldots, A_M)$ that takes values in $\mathcal{A}$, and hence $b \leq \mathbb{E}\{\psi(A_1, \ldots, A_M)\}$.

satisfies Assumption A3, and $\gamma'(\tau)$ is given by (85). Else (with probability $1-p$), for all slots $\tau \in \{rT, \ldots, rT+T-1\}$ choose $\alpha^*(t) = \alpha''_{\omega(\tau)}$ and $\gamma^*(\tau) = \gamma''$, where the decisions $\alpha''_{\omega(\tau)}$ and γ'' solve (15) and yield cost F_r^*.[12] Note from our construction here that either all slots of the frame use the first policy (which happens with probability p), or all slots of the frame use the second. Considering the expectation of the right-hand-side under this randomized policy, we have:

$$\Delta_T(rT) + \sum_{\tau=rT}^{rT+T-1} [V\hat{y}_0(\alpha(\tau), \omega(\tau)) + Vf(\gamma(\tau))] \leq$$
$$BT + CT + \tilde{D}T(T-1) + VT\epsilon_V$$
$$+(1-p)VTF_r^* + pT[Vy_0^{max} + Vf^{max}]$$
$$+ \sum_{l=1}^{L} Z_l(rT) \sum_{\tau=rT}^{rT+T-1} \left[\epsilon_Z - p(\delta - (\epsilon_H + \theta)\sum_{m=1}^{M}\beta_{l,m})\right]$$
$$+ \sum_{k=1}^{K} Q_k(rT) \sum_{\tau=rT}^{rT+T-1} [\epsilon_Q - p\delta]$$
$$+ \sum_{m=1}^{M} |H_m(rT)| \sum_{\tau=rT}^{rT+T-1} [\epsilon_H - p(\epsilon_H + \theta)]$$

Now choose the probability p to make all of the above queueing terms non-positive, as follows:

$$p \triangleq \max\left[\frac{\epsilon_Z}{\delta - (\epsilon_H + \theta)\beta_{sum}}, \frac{\epsilon_Q}{\delta}, \frac{\epsilon_H}{\epsilon_H + \theta}\right]$$

where

$$\beta_{sum} \triangleq \max_{l \in \{1,\ldots,L\}} \sum_{m=1}^{M} \beta_{l,m}$$

This is a valid probability (so that $0 \leq p \leq 1$) by definition of θ (being the solution of (83)). Therefore:

$$\Delta_T(rT) + \sum_{\tau=rT}^{rT+T-1} [V\hat{y}_0(\alpha(\tau), \omega(\tau)) + Vf(\gamma(\tau))] \leq$$
$$BT + CT + \tilde{D}T(T-1) + VT\epsilon_V$$
$$+(1-p)VTF_r^* + pTV[y_0^{max} + f^{max}]$$

Summing the above over $r \in \{0, 1, \ldots, R-1\}$, using non-negativity of the Lyapunov function, all queues are initially empty, convexity of $f(\cdot)$, and dividing by RTV yields:

$$\overline{y}_0 + f(\overline{\gamma}) \leq (1-p)\frac{1}{R}\sum_{r=0}^{R-1} F_r^* + \epsilon_V + p(y_0^{max} + f^{max})$$
$$+ \frac{(B + C + \tilde{D}(T-1))}{V}$$

---

[12]For simplicity, we assume here that the optimum of problem (15) is achievable by a single policy, else just take a policy that comes within ϵ of F_r^* and let $\epsilon \to 0$.

Finally, we have:

$$f(\overline{x}) \leq f(\overline{\gamma}) + \sum_{m=1}^{M} \nu_m |\overline{x}_m - \overline{\gamma}_m|$$
$$\leq f(\overline{\gamma}) + \sum_{m=1}^{M} \frac{\nu_m H_m^{max}}{RT}$$
$$\leq f(\overline{\gamma}) + \sum_{m=1}^{M} \frac{\nu_m V C_3}{\theta RT}$$

This proves Theorem 2c. $\qquad\square$

APPENDIX E — THE B AND D CONSTANTS

Values of B and D that satisfy (34) and (41) are given by:

$$B = \frac{1}{2}\sum_{l=1}^{L}(z_l^{diff})^2 + \frac{1}{2}\sum_{m=1}^{M}(h_m^{diff})^2$$
$$+ \frac{1}{2}\sum_{k=1}^{K}[(b_k^{max})^2 + (a_k^{max})^2] \qquad (90)$$
$$D = \frac{1}{2}\sum_{l=1}^{L}(z_l^{diff})^2 + \frac{1}{2}\sum_{m=1}^{M}(h_m^{diff})^2$$
$$+ \frac{1}{2}\sum_{k=1}^{K}(q_k^{diff})^2 \qquad (91)$$

where constants z_l^{diff}, h_m^{diff}, q_k^{diff} are defined after equation (41).

The Internet Model: For the internet model of Section IV, there are no queues $Q_k(t)$ and so we have $q_k^{diff} = a_k^{max} = b_k^{max} = 0$. We further have $h_m^{diff} = A_m^{max}$, and

$$z_l^{diff} = \max\left[C_l^{max}, \sum_{m=1}^{M} 1_{l,m} A_m^{max}\right]$$

where $1_{l,m}$ is equal to 1 if it is possible for session m to ever be routed over link l, and zero else. Using this value of z_l^{diff} in (90) and (91), the values of B and D for this internet context, and in particular for the utility bound (52), are:

$$B = D = \frac{1}{2}\sum_{l=1}^{L}(z_l^{diff})^2 + \frac{1}{2}\sum_{m=1}^{M}(A_m^{max})^2$$

The Dynamic Queueing Network Model: For the dynamic queueing network of Section V, there are no $Z_l(t)$ queues and so $z_l^{diff} = 0$. Further, $h_m^{diff} = A_m^{max}$. Because indices k of queues $Q_k(t)$ in the general framework correspond to indices (n, c) for queues $Q_n^{(c)}(t)$ in the dynamic queueing network, the values of B and D satisfy (34) and (41) whenever the following holds for all t:

$$B \geq \frac{1}{2}\sum_{m=1}^{M}(A_m^{max})^2 + \frac{1}{2}\sum_{n,c}[b_n^{(c)}(t)^2 + a_n^{(c)}(t)^2] \quad (92)$$
$$D \geq \frac{1}{2}\sum_{n,c}\max[b_n^{(c),max}, a_n^{(c),max}]\max[b_n^{(c)}(t), a_n^{(c)}(t)]$$
$$+ \frac{1}{2}\sum_{m=1}^{M}(A_m^{max})^2 \qquad (93)$$

We use this form, rather than the more explicit form (90), (91), because this often allows a tighter bound when we incorporate the structure of the network. Specifically, define constants $\mu_n^{max,in}$, $\mu_n^{max,sum}$, $x_n^{(c),max}$ as the maximum possible sum transmission rate into node n, sum of transmission rates into and out of node n, and exogenous arrivals to source node n of commodity c, respectively, over one slot. Note that $x_n^{(c),max}$ is given by:

$$x_n^{(c),max} = \sum_{m \in \mathcal{M}_n^{(c)}} A_m^{max}$$

Then we have for each $n \in \{1, \dots, N\}$:

$$\sum_c [b_n^{(c)}(t)^2 + a_n^{(c)}(t)^2]$$

$$\leq \sum_c \left[b_n^{(c)}(t)^2 + \left(\sum_{i=1}^N \mu_{in}^{(c)}(t) + x_n^{(c),max} \right)^2 \right]$$

$$= \sum_c \left[b_n^{(c)}(t)^2 + \left(\sum_{i=1}^N \mu_{in}^{(c)}(t) \right)^2 \right]$$

$$+ \sum_c \left[(x_n^{(c),max})^2 + 2 \left(\sum_{i=1}^N \mu_{in}^{(c)}(t) \right) (x_n^{(c),max}) \right]$$

$$\leq \left(\sum_c \left[b_n^{(c)}(t) + \left(\sum_{i=1}^N \mu_{in}^{(c)}(t) \right) \right] \right)^2$$

$$+ \sum_c (x_n^{(c),max})^2 + 2\mu_n^{max,in} \max_{c \in \{1,\dots,N\}} [x_n^{(c),max}]$$

$$\leq (\mu_n^{max,sum})^2$$

$$+ \sum_c (x_n^{(c),max})^2 + 2\mu_n^{max,in} \max_{c \in \{1,\dots,N\}} [x_n^{(c),max}]$$

Therefore a value of B that satisfies (92) is given by:

$$B = \frac{1}{2} \sum_{n=1}^N [(\mu_n^{max,sum})^2 + \sum_c (x_n^{(c),max})^2]$$

$$+ \sum_{n=1}^N \mu_n^{max,in} \max_{c \in \{1,\dots,N\}} [x_n^{(c),max}]$$

$$+ \frac{1}{2} \sum_{m=1}^M (A_m^{max})^2$$

Finally, define e_n as follows:

$$e_n \stackrel{\Delta}{=} \max_{c \in \{1,\dots,N\}} \max[b_n^{(c),max}, a_n^{(c),max}]$$

Then for each $n \in \{1, \dots, N\}$ we have:

$$\sum_c \max[b_n^{(c),max}, a_n^{(c),max}] \max[b_n^{(c)}(t), a_n^{(c)}(t)]$$

$$\leq e_n \sum_c \max[b_n^{(c)}(t), a_n^{(c)}(t)]$$

$$\leq e_n \sum_c [b_n^{(c)}(t) + a_n^{(c)}(t)]$$

$$\leq e_n [\mu_n^{max,sum} + \sum_c x_n^{(c),max}]$$

Therefore a value of D that satisfies (93) is:

$$D = \frac{1}{2} \sum_{m=1}^M (A_m^{max})^2$$

$$+ \frac{1}{2} \sum_{n=1}^N e_n [\mu_n^{max,sum} + \sum_c x_n^{(c),max}]$$

For example, consider a wireless network where data is measured in integer units of packets (assumed to have a fixed length). Suppose that at most one packet can be transmitted or received per node per slot, and that a packet cannot be transmitted and received on the same slot at the same node. Then we have $\mu_n^{max,sum} = \mu_n^{max,in} = 1$. Further, suppose there is at most one source at any given node (so that $M \leq N$), and no source can admit more than 1 packet per slot. Then $\sum_c x_n^{(c),max} = 1$ if node n is a source, and zero else, and $e_n = 2$ if node n is a source, and 1 else. There are M source nodes and $N - M$ non-source nodes, and so B and D are:

$$B = D = (N + 4M)/2$$

REFERENCES

[1] L. Georgiadis, M. J. Neely, and L. Tassiulas. Resource allocation and cross-layer control in wireless networks. *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1-149, 2006.

[2] M. J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. *Proc. IEEE INFOCOM*, March 2005.

[3] M. J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. *IEEE/ACM Transactions on Networking*, vol. 16, no. 2, pp. 396-409, April 2008.

[4] M. J. Neely. *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD thesis, Massachusetts Institute of Technology, LIDS, 2003.

[5] L. Tassiulas. Scheduling and performance limits of networks with constantly changing topology. *IEEE Trans. on Inf. Theory*, May 1997.

[6] M. J. Neely, E. Modiano, and C. E Rohrs. Dynamic power allocation and routing for time varying wireless networks. *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 89-103, January 2005.

[7] M. J. Neely. Delay analysis for maximal scheduling with flow control in wireless networks with bursty traffic. *IEEE Transactions on Networking*, vol. 17, no. 4, pp. 1146-1159, August 2009.

[8] M. J. Neely, E. Modiano, and C. E. Rohrs. Tradeoffs in delay guarantees and computation complexity for $n \times n$ packet switches. *Proc. of Conf. on Information Sciences and Systems (CISS), Princeton*, March 2002.

[9] M. J. Neely and R. Urgaonkar. Cross layer adaptive control for wireless mesh networks. *Ad Hoc Networks (Elsevier)*, vol. 5, no. 6, pp. 719-743, August 2007.

[10] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener. Rethinking information theory for mobile ad hoc networks. *IEEE Communications Magazine*, vol. 46, no. 12, pp. 94-101, Dec. 2008.

[11] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory*, IT-23, no. 3, pp. 337-343, May 1977.

[12] T. M. Cover. Universal portfolios. *Mathematical Finance*, vol. 1, no. 1, pp. 1-29, Jan. 1991.

[13] T. M. Cover and E. Ordentlich. Universal portfolios with side information. *IEEE Transactions on Information Theory*, vol. 42, no. 2, 1996.

[14] N. Merhav and M. Feder. Universal schemes for sequential decision from individual data sequences. *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1280-1292, July 1993.

[15] M. J. Neely. Stock market trading via stochastic network optimization. *ArXiv Technical Report*, arXiv:0909.3891v1, Sept. 2009.

[16] J. A. Garay and I. S. Gopal. Call preemption in communication networks. *Proc. IEEE INFOCOM*, vol. 44, pp. 1043-1050, Florence, Italy, 1992.

[17] S. Plotkin. Competitive routing of virtual circuits in atm networks. *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 1128-1136, Aug. 1995.

[18] L. Lin, N. B. Shroff, and R. Srikant. Asymptotically optimal power-aware routing for multihop wireless networks with renewable energy sources. *Proc. IEEE INFOCOM*, March 2005.

[19] J. J. Jaramillo and R. Srikant. Admission control and routing in multi-hop wireless networks. *Proc. IEEE Conf. on Decision and Control (CDC)*, Dec. 2008.

[20] M. Andrews. Maximizing profit in overloaded networks. *Proc. IEEE INFOCOM*, March 2005.

[21] A. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Systems*, vol. 50, pp. 401-457, 2005.

[22] A. Stolyar. Greedy primal-dual algorithm for dynamic resource allocation in complex networks. *Queueing Systems*, vol. 54, pp. 203-220, 2006.

[23] F.P. Kelly, A.Maulloo, and D. Tan. Rate control for communication networks: Shadow prices, proportional fairness, and stability. *Journ. of the Operational Res. Society*, 49, p.237-252, 1998.

[24] S. H. Low and D. E. Lapsley. Optimization flow control, i: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, vol. 7(6): 861-75, Dec. 1999.

[25] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle. Layering as optimization decomposition: A mathematical theory of network architectures. *Proceedings of the IEEE*, vol. 95, no. 1, Jan. 2007.

[26] E. Leonardi, M. Mellia, M. A. Marsan, and F. Neri. Optimal scheduling and routing for maximizing network throughput. *IEEE/ACM Transactions on Networking*, vol. 15, no. 6, Dec. 2007.

[27] Y. Li, A. Papachristodoulou, and M. Chiang. Stability of congestion control schemes with delay sensitive traffic. *Proc. IEEE ACC, Seattle, WA*, June 2008.

[28] J. K. MacKie-Mason and H. R. Varian. Pricing congestible network resources. *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, September 1995.

[29] D. P. Bertsekas and R. Gallager. *Data Networks*. New Jersey: Prentice-Hall, Inc., 1992.

[30] M. J. Neely and E. Modiano. Convexity in queues with general inputs. *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 706-714, Feb. 2005.

[31] X. Lin and N. B. Shroff. Joint rate control and scheduling in multihop wireless networks. *Proc. of 43rd IEEE Conf. on Decision and Control, Paradise Island, Bahamas*, Dec. 2004.

[32] M. J. Neely. Energy optimal control for time varying wireless networks. *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 2915-2934, July 2006.

[33] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1995.

[34] M. J. Neely. Delay-based network utility maximization. *Proc. IEEE INFOCOM*, 2010.